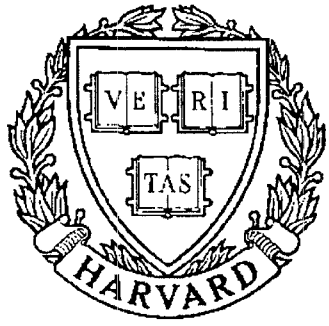
TECHNICAL RESEARCH REPORT

SYSTEMS RESEARCH CENTER



Supported by the National Science Foundation Engineering Research Center Program (NSFD CD 8803012), the University of Maryland, Harvard University, and Industry

Research support for this report has been partially provided by the Office of Naval Research under Contract N00014-89-J-1375

Quantization and Fusion for Multi-Sensor Discrimination from Dependent Observations

by Y.A. Chau and E. Geraniotis

TR 91-27

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1991		**00-00-1991 to 00-00-1991**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Quantization and Fusion for Multi-Sensor Discrimination from Dependent Observations	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Maryland,Systems Research Center,College Park,MD,20742**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**70**	

QUANTIZATION AND FUSION FOR MULTI-SENSOR DISCRIMINATION FROM DEPENDENT OBSERVATIONS

Yawgeng A. Chau and Evaggelos Geraniotis

QUANTIZATION AND FUSION FOR MULTI-SENSOR DISCRIMINATION FROM DEPENDENT OBSERVATIONS

Yawgeng A. Chau and Evaggelos Geraniotis

Department of Electrical Engineering
and Systems Research Center
University of Maryland
College Park, MD 20742

ABSTRACT

Schemes for quantization and fusion in multi-sensor systems used for discriminating between two sequences of dependent observations are introduced and analyzed. The observation sequences of each sensor under the two hypotheses are arbitrary stationary dependent sequences that can not be modeled as signal in additive noise; the objective of the fusion center is to disrciminate between the two hypotheses. These observation models are well motivated by practical multi-sensor target discrimination problems. Two cases are considered: in the first, the observation sequences of the sensors are individually dependent but jointly mutually independent; in the second case, the observation sequences are dependent across both time and sensors. The dependence in the observations across time and/or sensors is modeled by m-dependent, ϕ-mixing, or ρ-mixing processes. The following four quantization/fusion schemes are considered: (a) forming test statistics at the sensors by passing the observations through memoryless nonlinearities, summing them up, and fusing these test statistics without previous quantization; (b) quantizing uniformally (with equidistant breakpoints) each sensor observation and then fusing; (c) quantizing optimally each sensor observation and then fusing; and (d) using the sensor test statistic of (a) to make binary decisions and then fusing the binary decisions. To guarantee high-quality performance, a common large sample size is employed by each sensor and an asymptotic analysis is pursued. Design criteria are developed from the Bayesian cost of the fusion center for deriving the optimal memoryless nonlinearities of the sensor test statistics and the sensor quantizer parameters (quantization levels and breakpoints). These design criteria are shown to involve an extension of the generalized signal-to-noise ratio used in single-sensor detection and quantization. The optimal nonlinearities and quantizers are obtained as the solutions of linear coupled or uncoupled integral equations involving the univariate and bivariate probability densities of the sensor observations. Numerical results based on simulation are presented for specific cases of practical interest to compare the relative performance of the four quantization/fusion schemes described above and to establish their superiority to schemes that ignore the dependence across time and/or sensors in the observations.

This research was supported in part by the Office of Naval Research under contract N00014-89-J-1375 and in part by the Systems Research Center at the University of Maryland, College Park, through the National Science Foundation's Engineering Research Centers Program: NSF CDR 8803012.

I. INTRODUCTION AND PROBLEM FORMULATION

In most of the recent works on multi-sensor detection with a fusion center (see [1]-[5]), the observations are assumed to be independent across time (by being modeled as i.i.d. sequences of random variables) and sensors (through the assumption that the observation sequences of the different sensors are mutually independent, when conditioned on a particular hypothesis being true), both of which are intended to make the analysis tractable. Yet the observations are generally dependent. Indeed, the observation processes of the sensors become dependent across time for each sensor, when the sampling rates increase, and correlated across sensors, when the locations of the sensors are close geographically.

In [6], the problem of distributed detection of a signal in continuous-time correlated additive Gaussian noise is addressed and the optimal decision test for each sensor derived, for a fixed fusion rule; the results are limited to Gaussian noise statistics and a signal-in-additive-noise configuration. In [7], a fusion scheme with a two-bit fuzzy decision at each sensor is considered for the case of a single observation with dependence across sensors; this is an interesting formulation but not a sufficiently general one. In our work of [8], we treat multi-sensor detection problems with correlated observations across time and/or sensors of general m-dependent or mixing type. However, although the decision rules of the sensors were coupled through the optimization of a common cost function, no quantization or fusion of the sensor decisions or observations was considered in [8]. In our other recent work of [9] (a companion paper), quantization and fusion schemes were considered in a multi-sensor context for the purpose of detecting the presence of weak signals in stationary dependent noise.

In this paper, we consider **quantization and fusion in multi-sensor systems for the discrimination between arbitrary observation sequences under the two hypotheses which are dependent across time and/or sensors.** This work complements that of [9] and extends it significantly by considering non-weak-signal models for the observations of the sensors. In particular, the observations sequence of each sensor under the two hypotheses are arbitrary (not necessarily signals in additive noise) stationary dependent sequences. The weak-signal model of [9] and the associated locally optimum detection approach are not applicable here; a **nonlocal detection approach** is adopted. For all the quantization and fusion schemes of interest to this paper, we consider two operational scenarios, one in which the observation sequences of the various sensors under each hypothesis are characterized by dependence across

time only and another characterized by dependence across both time and sensors.

In our model, the sensors collect n observations each (n being the common sample size) and either form test statistics, which, after quantization, are transmitted to the fusion center, or directly transmit to the fusion center a quantized version of their observations. At the fusion center, the information sent by the sensors is processed further and a decision is reached as to which of two hypotheses, H_0 or H_1, is true. Here we introduce and analyze four quantization and fusion schemes. In all four schemes, the sensors employ memoryless nonlinearities, while the fusion center conducts an appropriate likelihood ratio test (and a different one for each scheme) with a threshold that can be easily determined. These schemes are described in detail later in this section.

To model the dependence in the observations of each individual sensor and between observations of different sensors we use **stationary mixing processes**, such as stationary m-dependent, ϕ-mixing, or ρ-mixing processes. The definitions of these mixing processes and the associated central limit theorems are detailed in [10] and [11]. These models of dependence have been successfully used in single-sensor discrimination problems (see [12]-[15]).

As a rule, the optimal detection schemes under the above models of dependence, for either single-sensor or multi-sensor/fusion-center configurations, involve high-order (larger than two) probability densities of the observations, which not only are difficult to characterize, but also impose prohibitive requirements on processor memory and complexity for the storage and processing of the dependent data. To avoid these complications in our analysis, design, and implementation, we employ **suboptimal sensor decision statistics, and sensor quantization schemes based on memoryless nonlinearities**; the fusion rules employed at the fusion center are basically likelihood ratio tests based on these suboptimal sensor decision statistics or quantization schemes. These are easier to implement and their derivation requires knowledge of only the univariate and bivariate probability densities of the observation sequences. Large sample sizes are necessary for the validity of the analysis and the usefulness of the resulting detectors; but these are well suited to hypothesis testing problems, for which a high-quality performance is desirable involving the ability to discriminate between two sequences of observations with very similar statistical description. Our analysis, which is based on a nonlocal detection approach, is valid under large sample sizes for any statistical relationship between the observation sequences under the two hypotheses; yet it is the case involving close similarity in the statistical descrip-

tions, for which our approach is most appropriate, as high-quality discrimination necessitates large sample sizes. For single-sensor systems, these memoryless nonlinearities were successfully used in [12]-[13] for nonlocal detection (discrimination) problems with m-dependent or mixing stationary observations, and in [14]-[15] for designing optimal quantizers, by considering the quantizers as special nonlinearities.

We now describe the four schemes (termed Schemes 1 to 4) for quantization and fusion in multi-sensor detection systems with dependent observations, whose analysis, optimal design, and performance evaluation (via simulation) constitute the subject of this paper. In our description, we focus attention on a two-sensor configuration; the extension to configurations with more sensors is discussed in a later section. First we introduce the notation for the model of the sensor observation sequences, the memoryless nonlinearities, the test statistics used by the sensors, and characterize the necessary probability distributions for the statistical description of our multi-sensor/fusion-center system.

The distributed discrimination problem considered in this paper is formulated as the following distributed binary hypotheses testing problem:

$$\begin{aligned} H_0^{(k)} &: X_i{}^{(k)} \sim f_{k,0}^{(n)}(\underline{X}^{(k)}) \\ H_1^{(k)} &: X_i{}^{(k)} \sim f_{k,1}^{(n)}(\underline{X}^{(k)}); \quad k = 1,2 \end{aligned} \tag{1}$$

where $\underline{X}^{(k)} = (X_1^{(k)}, X_2^{(k)}, \cdots, X_n^{(k)})$ is the vector of n stationary observations of sensor k and $f_{k,i}^{(n)}$ the associated n-th order multivariate density of the observation process for sensor k under $H_i^{(k)}$ $(i = 0, 1)$. The above model of the sensor observations is motivated by practical problems of naval discrimination between targets (ships) and decoys (chaff or active decoys). As is shown below, only the univariate (marginal) and bivariate (second-order joint) pdfs of the sensor observation sequences are involved in our analysis. The observations sequences are not modeled as signal in additive noise. As mentioned already, the two observation sequences $\{X_l^{(k)}\}_{l=1}^n$ $(k = 1, 2)$ are modeled by stationary m-dependent, or mixing-type processes such as ϕ-mixing or ρ-mixing processes. As discussed above, we are interested in worst-case situations, in which the statistical description of the observation sequences under the two hypotheses are similar, because large sample sizes are then necessary, if high-quality discrimination is desirable. For example, if $E_1[X_1^{(k)}] = E_0[X_1^{(k)}]$ and $E_1[(X_1^{(k)})^2] = E_0[(X_1^{(k)})^2]$ $(k = 1, 2)$ (i.e. the means and the powers of the observations under the two hypotheses are equal), a large sample size is required to achieve

small error probabilities at the fusion center.

Let $g_k(\cdot)$ be the memoryless nonlinearity for sensor k. In general, the argument of g_k is a continuous-amplitude (real) variable and g_k can take all real values. (In reality g_k is implemented through a discretized form with a large but finite number of amplitude levels, a matter to which we will return later). When quantization is employed, $g_k(\cdot)$ takes the particular form of a quantizer characterized by a finite number quantization levels and breakpoints. The test statistic for each sensor k basically has the form

$$T_{n,k} = \frac{1}{n}\sum_{l=1}^{n} g_k(X_l^{(k)}); \quad k = 1,2 \tag{2}$$

where n is the common (large) sample size, and $\{X_l^{(k)}\}_{l=1}^n$ a stationary observation sequence characterized by m-dependent or mixing type dependence. These mixing processes are described in detail in the tutorial [10]; we will not repeat the definitions here.

In general, the true distribution of $T_{n,k}$ $(k = 1,2)$ is difficult to obtain. However, for large n, one can employ the central limit theorem for the aforementioned mixing processes (see [11]). Specifically, under hypothesis $H_i^{(k)}$ $(k = 1,2; i = 0,1)$, define

$$\mu_{k,i}(g_k) = E_i[T_{n,k}]; \quad k = 1,2 \tag{3}$$

$$\sigma_{k,i}^2(g_k) = var_i[g_k(X_1^{(k)})] + 2\sum_{j=1}^{m_{k,i}} cov_i[g_k(X_1^{(k)}), g_k(X_{j+1}^{(k)})] \tag{4}$$

where $var_i[\cdot]$ and $cov_i[\cdot]$ are the variance and the covariance operators, respectively, and

$$\rho_i(g_1,g_2) = \frac{E_i[(T_{n,1}-\mu_{1,i})\cdot(T_{n,2}-\mu_{2,i})]}{(\frac{1}{n}\sigma_{1,i}\sigma_{2,i})} \tag{5}$$

where the numerator of (5) is equal to $\frac{1}{n}[cov_i\{g_1(X_1^{(1)}), g_2(X_1^{(2)})\} + 2\sum_{j=1}^{m_{1,2,i}} cov_i\{g_1(X_1^{(1)}), g_2(X_{j+1}^{(2)})\}]$; $m_{k,i}$ $(k = 1,2)$ is the range of dependence under $H_i^{(k)}$, for the m-dependent observations of the k-th sensor; and $m_{1,2,i}$ is the corresponding parameter for the cross-dependence of the observations of sensors 1 and 2. For mixing-type observations we take $m_{l,i} \to \infty$, for $l = 1,2$ and $(1,2)$. Clearly, under hypothesis $H_i^{(k)}$, $\mu_{k,i}(g_k)$ depends on the univariate pdf of the observation sequence $\{X_l^{(k)}\}_{l=1}^n$ denoted by $f_{k,i}(\cdot)$; $\sigma_{k,i}^2(g_k)$ depends on $f_{k,i}$ and the bivariate pdfs of $(X_1^{(k)}, X_{j+1}^{(k)})$, for $j = 1,2,\cdots$, denoted by $f_{k,i}^{(j)}(\cdot,\cdot)$; and $\rho_i(g_1,g_2)$ depends on $f_{1,i}$, $f_{2,i}$ and the bivariate pdfs of $(X_1^{(1)}, X_{j+1}^{(2)})$, for $j = 0,1,2,\cdots$, denoted by $f_{1,2,i}^{(j)}(\cdot,\cdot)$. The primary distinction between the above definitions for the multi-sensor discrimination problem treated in this paper

and the problem of multi-sensor weak-signal detection considered in [9] is that the variances under the two hypotheses are generally different in this paper, whereas they are the same in [9].

According to the central limit theorem for mixing processes, under $H_i^{(k)}$ $(i = 0, 1)$ and suitable conditions (see [10]-[11]), if $\sigma_{k,i}^2$ converges (which is trivially satisfied for m-dependent processes but needs to be assumed for mixing-type processes) and $\sigma_{k,i}^2 > 0$, $(T_{n,k} - \mu_{k,i})/(\frac{1}{\sqrt{n}}\sigma_{k,i})$ is asymptotically Gaussian distributed with the standard $\mathcal{N}(0, 1)$ pdf. The conditions for the validity of the central limit theorem for each individual type of mixing processes are described in [11] and are assumed to be satisfied in this paper. For example, ρ-mixing processes (for which the numerical results in this paper are generated) require not only the convergence of the infinite series above and the strict positivity of the variance, but also the convergence of the series $\sum_{n=1}^{\infty} \rho_{i,2^n}^{(l)}$, where $\rho_{i,n}$ denotes the sequence of ρ-mixing parameters under hypothesis $H_i^{(k)}$; the superscript (l) is (1) for the observations sequence of sensor 1, (2) for that of sensor 2, and (1, 2) for the jointly considered observation sequences of sensors 1 and 2.

Henceforth, we suppose that, under hypothesis $H_i^{(k)}$ $(i = 0, 1)$, $T_{n,k}$ of sensor k is asymptotically Gaussian distributed with mean $\mu_{k,i}(g_k)$ and variance $\sigma_{k,i}^2(g_k)/n$. Moreover, the pair $(T_{n,1}, T_{n,2})$ is asymptotically jointly Gaussian distributed with the correlation coefficient $\rho_i(g_1, g_2)$ $(-1 \leq \rho_i(g_1, g_2) \leq 1)$. For the sake of notational convenience, we omit the arguments of the means $\mu_{k,i}$, the variances $\sigma_{k,i}$, and the correlation coefficients ρ_i. We also use g_k and $T_{n,k}$ to represent the nonlinearity or quantizer and the associated test statistic, respectively, in the general expressions. We then use different notation to emphasize their function in the different quantization and fusion schemes considered in this paper. Similar notation rules are applied to the means, variances, and correlation coefficients.

The four quantization and fusion schemes central to this paper are defined as follows: In **Scheme 1** (illustrated in Fig. C.1), the k-th sensor employs the nonlinearity g_k, forms the test statistic $T_{n,k}$ given by (2), and transmits it directly to the fusion center, where a likelihood ratio test is performed based on $T_{n,k}$, for $k = 1, 2$. If one could reliably transmit a real number through a bandlimited channel (as are the channels between the sensors and the fusion center), this fusion scheme, which is optimal within the class of schemes employing memoryless nonlinearities in the sensor test statistics, would also be practical. However, in reality we can only transmit a finite number of bits of information through the aforementioned channels. Therefore, quantization of the test statistics or the sensor observations themselves is of interest and this is considered by

the following three schemes.

In **Scheme 2** (illustrated in Fig. C.2), the observation $X_l^{(k)}$ of sensor k at time l is first quantized by a uniform quantizer $\bar{g}_k$, which is obtained from the discrete-form of the optimal nonlinearity g_k of Scheme 1, and then transmitted to the fusion center, where the test statistic

$$\bar{T}_{n,k} = \frac{1}{n}\sum_{l=1}^{n} \bar{g}_k(X_l^{(k)}) \tag{6}$$

is formed. To distinguish the special nonlinearity (a quantizer) and the associated statistic of Scheme 2 from those of Scheme 1, we use $\bar{g}_k$ and $\bar{T}_{n,k}$ for representing them. Finally, the fusion center performs a likelihood ratio test based on $\bar{T}_{n,k}$, for $k = 1, 2$. This scheme is motivated by the need to reduce the information that the sensors transmit to the fusion center. It is simple to implement, because its breakpoints are uniformly spaced over the interval defining the support of the sensor observations; for a large number of quantization levels it is supposed to approximate Scheme 1.

Scheme 3 is similar to Scheme 2, except that the quantizers are now optimized and can be obtained from the discrete-form of the nonlinearities of Scheme 1. In order to distinguish this scheme from Schemes 1 and 2, the optimal quantizer (breakpoints and quantization levels) and the associated test statistic of Scheme 3 are denoted by Q_k and $\tilde{T}_{n,k}$, respectively, where

$$\tilde{T}_{n,k} = \frac{1}{n}\sum_{l=1}^{n} Q_k(X_l^{(k)}). \tag{7}$$

This particular scheme is motivated by the need to approximate Scheme 1 by a quantization method for the usual reasons of reduction in the transmitted information, but it is also supposed to achieve better performance than that of Scheme 2 with fewer quantization levels (and thus transmitted bits of information), because of the optimized break-points and levels.

Finally, in **Scheme 4** (illustrated in Fig. C.3), the sensors form the test statistics $T_{n,k}$ as in Scheme 1, then use thresholds to make hard binary decisions about which hypothesis (H_1 or H_0) is true, and then transmit these binary decisions to the fusion center. The fusion center then executes a likelihood ratio test based on the sensor decisions in order to reach the final decision. This scheme has the same configuration as Scheme 4. However, the number of quantization levels for this scheme is two; thus we can not use the analysis and optimal design developed for Scheme 4, which is valid only if the the number of quantization levels is large.

As described above, the function of the fusion center is to collect (as in Schemes 1 and 4) or form (as in Schemes 2 and 3) the test statistics from the information transmitted by the sensors. In the design of optimal nonlinearities or quantizers for Schemes 1-4, we employ the Bayesian risk criterion on the error probabilities of the fusion center. We use large deviations theory to associate the Bayes risk with the final design criterion for each nonlinearity or quantizer that has the form of **generalized signal-to-noise ratio.**

The course of obtaining the optimal nonlinearities or quantizers involves solving **linear coupled or uncoupled integral equations**, which depend on the **univariate and bivariate probability density functions** (pdfs) of the sensor noise sequence. The optimal threshold of each sensor η_k is determined by the optimization of the individual design criterion for each case; it depends only on the means and the variances of the test statistics $T_{n,k}$, under the two hypotheses, and thus on the aforementioned univariate and bivariate pdfs.

To recapitulate, the contributions of this paper are the following:
(i) it investigates asymptotically optimal memoryless schemes for quantizing either the sensor observations or the sensor test statistics, before transmitting to the fusion center, and appropriate fusion rules for multi-sensor/fusion-center discrimination problems with dependence in the observations across-time and sensors;
(ii) it extends the useful methodology of nonlocal memoryless detection (discrimination) from the single-sensor case to the case of multiple sensors with a fusion center:
(iii) it treats multi-sensor/fusion-center discrimination problems with dependence in the observations across-time and sensors, thus extending the existing work in this area which deals with i.i.d. sensor observations.

The remainder of this paper is organized as follows. In Section II, preliminary comments apply to all quantization/fusion schemes considered in this paper are provided. In Section III, our four quantization/fusion schemes are presented for the case of dependence across time only. Then the same quantization/fusion schemes are presented in Section IV for the case of dependence across time and sensors; the important special case of sensors with identical univariate and bivariate probability densities is also treated there. In Section V, numerical results for the two-sensor/fusion-center discrimination between lognormal and Rayleigh dependent observations sequences are provided. In Section VI, the robustification of our schemes to statistical uncertainty in the sensor observations, the extension of our schemes to environments with more

than two sensors, and the conclusions drawn from our work are discussed.

II. PRELIMINARIES

Since a large sample size is used, we are interested in the asymptotic exponential rate of decrease of the Bayesian cost defined by

$$E[C] = c_0 p P_0(\ln L_n > \eta) + c_1(1-p)P_1(\ln L_n \leq \eta) \tag{8}$$

where p is the a priori probability of H_0 and c_i $(i = 0, 1)$ are positive constants. When $c_i = 1$ $(i = 0, 1)$, this Bayesian cost is the error probability in the fusion center. The exponential rate of this expected cost is defined by $-\frac{1}{n}\ln E[C]$.

1. Lower Bound on the Asymptotic Rate of Error Probabilities for Schemes 1 - 3

Here we derive a lower bound on the asymptotic rate of $E[C]$ (or upper bound on $E[C]$), which is then used in the design of optimal nonlinearities or quantizers in Schemes 1-3.

Since $T_{n,1}, T_{n,2}$ are asymptotically jointly Gaussian, the log-likelihood ratio function of the fusion center takes the form

$$\begin{aligned}\ln L_n &= \frac{n}{2(1-\rho_0^2)}\left[\frac{(T_{n,1}-\mu_{1,0})^2}{\sigma_{1,0}^2} + \frac{(T_{n,2}-\mu_{2,0})^2}{\sigma_{2,0}^2} - \frac{2\rho_0(T_{n,1}-\mu_{1,0})(T_{n,2}-\mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right] \\ &\quad -\frac{n}{2(1-\rho_1^2)}\left[\frac{(T_{n,1}-\mu_{1,1})^2}{\sigma_{1,1}^2} + \frac{(T_{n,2}-\mu_{2,1})^2}{\sigma_{2,1}^2} - \frac{2\rho_1(T_{n,1}-\mu_{1,1})(T_{n,2}-\mu_{2,1})}{\sigma_{1,1}\sigma_{2,1}}\right] \\ &\quad +\ln\frac{\sigma_{1,0}\sigma_{2,0}}{\sigma_{1,1}\sigma_{2,1}}\end{aligned} \tag{9}$$

which is a quadratic and not a linear combination of $T_{n,1}$ and $T_{n,2}$, and thus not asymptotically Gaussian distributed.

The means of $\ln L_n$ under the two hypotheses are given by

$$\begin{aligned}E_0[\ln L_n] &= 1 + \ln\frac{\sigma_{1,0}\sigma_{2,0}}{\sigma_{1,1}\sigma_{2,1}} - \frac{\sigma_{1,0}^2}{2(1-\rho_1^2)\sigma_{1,1}^2} - \frac{\sigma_{2,0}^2}{2(1-\rho_1^2)\sigma_{2,1}^2} - \frac{\rho_0\rho_1\sigma_{1,0}\sigma_{2,0}}{(1-\rho_1^2)\sigma_{1,1}\sigma_{2,1}} \\ &\quad -\frac{n}{2(1-\rho_1^2)}\left[\frac{(\mu_{1,1}-\mu_{1,0})^2}{\sigma_{1,1}^2} + \frac{(\mu_{2,1}-\mu_{2,0})^2}{\sigma_{2,1}^2} - \frac{2\rho_1(\mu_{1,1}-\mu_{1,0})(\mu_{2,1}-\mu_{2,0})}{\sigma_{1,1}\sigma_{2,1}}\right]\end{aligned} \tag{10}$$

and

$$\begin{aligned}E_1[\ln L_n] &= -1 + \ln\frac{\sigma_{1,0}\sigma_{2,0}}{\sigma_{1,1}\sigma_{2,1}} + \frac{\sigma_{1,1}^2}{2(1-\rho_0^2)\sigma_{1,0}^2} + \frac{\sigma_{2,1}^2}{2(1-\rho_0^2)\sigma_{2,0}^2} + \frac{\rho_0\rho_1\sigma_{1,1}\sigma_{2,1}}{(1-\rho_0^2)\sigma_{1,0}\sigma_{2,0}} \\ &\quad +\frac{n}{2(1-\rho_0^2)}\left[\frac{(\mu_{1,1}-\mu_{1,0})^2}{\sigma_{1,0}^2} + \frac{(\mu_{2,1}-\mu_{2,0})^2}{\sigma_{2,0}^2} - \frac{2\rho_0(\mu_{1,1}-\mu_{1,0})(\mu_{2,1}-\mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right]\end{aligned} \tag{11}$$

moreover, as $n \to \infty$,

$$\frac{1}{n}E_0[\ln L_n] \to -\frac{1}{2(1-\rho_1^2)}\left[\frac{(\mu_{1,1}-\mu_{1,0})^2}{\sigma_{1,1}^2}+\frac{(\mu_{2,1}-\mu_{2,0})^2}{\sigma_{2,1}^2}-\frac{2\rho_1(\mu_{1,1}-\mu_{1,0})(\mu_{2,1}-\mu_{2,0})}{\sigma_{1,1}\sigma_{2,1}}\right]=\varphi_0 \tag{12}$$

and

$$\frac{1}{n}E_1[\ln L_n] \to \frac{1}{2(1-\rho_0^2)}\left[\frac{(\mu_{1,1}-\mu_{1,0})^2}{\sigma_{1,0}^2}+\frac{(\mu_{2,1}-\mu_{2,0})^2}{\sigma_{2,0}^2}-\frac{2\rho_0(\mu_{1,1}-\mu_{1,0})(\mu_{2,1}-\mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right]=\varphi_1. \tag{13}$$

In other words, $E_i[\ln L_n]$ $(i=0,1)$ are dominated by $n\cdot\varphi_i$, respectively, as $n \to \infty$. To guarantee the reasonable condition that the error probabilities converge to zero, as $n \to \infty$, we assume that the consistency condition described by

$$\varphi_0 \le \eta \le \varphi_1 \tag{14}$$

for $i=0,1$, holds for the discrimination problem of interest. From the ergodic mean theorem, the condition (14) guarantees that the error probability in the fusion center converges to zero, as the sample size increases. A consistency condition similar to (14) was used in the single-sensor environment, but with a linear test, meaning that the log-likelihood ratio $\ln L_n$ was a linear function of the test statistics $T_{n,k}$ (see [12] and [13]).

Let $P_{max} = \max\{P_0(\ln L_n > \eta), P_1(\ln L_n \le \eta)\}$. Then, as $n \to \infty$, we have

$$-\frac{1}{n}\ln E[C] = -\frac{1}{n}\ln P_{max} + -\frac{1}{n}\ln\left[c_0 p\frac{P_0(\ln L_n > \eta)}{P_{max}} + c_1(1-p)\frac{P_1(\ln L_n \le \eta)}{P_{max}}\right] \to -\frac{1}{n}\ln P_{max} \tag{15}$$

To establish the design criteria for the optimal sensors thresholds and nonlinearities or quantizers, we first apply the large deviation principle of [17] to characterize the exponential rates of the two types of error probabilities, and then derive lower bounds on these rates (or upper bounds on the error probabilities), which are then used in the design of optimal sensor nonlinearities or quantizers.

Now let us give some definitions associated with the large deviation principle, which are relevant to our treatment. For $s \in \mathrm{R}$, define

$$\bar{b}_i(s) = \lim_{n\to\infty} b_{i,n}(s) = \lim_{n\to\infty}\frac{1}{n}\ln E_i[\exp(s\ln L_n)] \tag{16}$$

$$\bar{D}_i(\bar{b}_i) = \{s \in \mathrm{R}^2 : \bar{b}_i(s) < \infty\} \tag{17}$$

and

$$\bar{I}_i(z) = \sup_{s \in \mathrm{R}} \{sz - \bar{b}_i(s)\} \tag{18}$$

where $\bar{I}_i$ $(i = 0, 1)$ are the Legendre-Fenchel transforms of $\bar{b}_i$. Moreover, under $H_i^{(k)}$ $(i = 0, 1)$ and for a subset G of R, define

$$\bar{I}_i(G) = \inf\{\bar{I}(z) : z \in G\} \tag{19}$$

and the induced probabilities

$$P_i(G) = Pr_i\{\ln L_n \in G\} \tag{20}$$

where $\bar{I}_i(G)$ are termed entropy functions in [17]. It is easy to show that $\bar{b}_{i,n}(s_1, s_2)$ $(i = 0, 1)$ are convex functions, for all n, hence $\bar{b}_i(s_1, s_2)$ are also convex functions. The following theorem characterizes the asymptotic rates of the two types of error probabilities and is a direct result of Theorem II.2 in [17].

Theorem 1: Under $H_i^{(k)}$ $(i = 0, 1)$, suppose $\bar{b}_i(s)$ exist for all $s \in \mathrm{R}$ as closed functions, and that $\bar{D}_i(\bar{b}_i)$ has non-empty interior containing the point $s = 0$. Then, for any closed subset F of R,

$$\liminf_{n \to \infty} -\frac{1}{n} \ln P_i(F) \geq \bar{I}_i(F). \tag{21}$$

Furthermore, if $\bar{b}_i$ is differentiable on the interior of $\bar{D}_i$ and steep in the sense that the magnitude of the derivative of $\bar{b}_i$, $|d\bar{b}_i(s)/ds|$ diverges to infinity, as s tends to the boundary of $\bar{D}_i$, then for any open subset E of R,

$$\limsup_{n \to \infty} -\frac{1}{n} \ln P_i(E) \leq \bar{I}_i(E), \tag{22}$$

In general, under the true distribution of $H_i^{(k)}$, the exact form of $\bar{I}_i$ is difficulty to obtain, especially for the quadratic test characterized above. Even by assuming the joint Gaussian distribution for $(T_{n,1}, T_{n,2})$, we cannot have useful criteria for optimal nonlinearities or quantizers, because the supreme of $sz - b_i(s)$ in (18) does not take a neat form. Alternatively, let us consider the local behavior of $sz - \bar{b}_i(s)$ for $|s| < \bar{s}$ with $\bar{s} \to 0$. Let the set B_s be

$$B_s = \{s \in \mathrm{R} : |s| < \bar{s}; \bar{s} \to 0\} \tag{23}$$

which is a subset of **R**. Then we have

$$\bar{I}_i(z) = \sup_{s \in \mathrm{R}} \{sz - \bar{b}_i(s)\} \geq \sup_{s \in B_s} \{sz - \bar{b}_i(s)\} = I_i(z) \tag{24}$$

where $I_i(z)$ $(i = 0, 1)$ are characterized by the following lemma, whose proof, together with the proof of Theorem 2 that follows, are given in Appendix A.

Lemma 1: The lower bounds $I_i(z)$ $(i = 0, 1)$ on the asymptotic rates $\bar{I}_i(z)$ are given by

$$I_i(z) = (z - \varphi_i)\tilde{s}_i \tag{25}$$

where $\tilde{s}_i = \arg\sup_{s \in B_s}\{sz - \bar{b}_i(z)\}$ has a sign, which makes I_i positive for any z in the particular set G, and φ_i are defined by (12), for $i = 0$, and (13), for $i = 1$.

Thus, for any set G, we have

$$\bar{I}_i(G) = \inf_{z \in G} \bar{I}_i(z) \geq \inf_{z \in G} I_i(z) = I_i(G) \tag{26}$$

where $I_i(G)$ is defined in a way similar to $\bar{I}_i(G)$.

In other words, as $n \to \infty$, if $-\frac{1}{n}P_i(G) \to \bar{I}_i(G)$, then $P_i(G) \leq \exp\{-nI_i(G)\}$ by using (26). According to this lemma, we have the following theorem, which characterizes the upper bounds on the two types of error probabilities of the fusion center under a large sample size.

Theorem 2: Suppose the conditions in Theorem 1 for (21) and (22) are satisfied. Moreover, suppose that, under $H_i^{(k)}$ $(i = 0, 1; k = 1, 2)$, $T_{n,k}$ are asymptotically Gaussian distributed with mean $\mu_{k,i}$ and variance $\sigma_{k,i}^2$ and $(T_{n,1}, T_{n,2})$ asymptotically jointly Gaussian distributed with the correlation coefficient ρ_i. Then, for $\ln L_n$ given by (9), as $n \to \infty$,

$$-\frac{1}{n}\ln P_0(\ln L_n > n\eta) \geq (\eta - \varphi_0)\tilde{s}; \quad -\frac{1}{n}\ln P_1(\ln L_n \leq n\eta) \geq (\varphi_1 - \eta)\tilde{s} \tag{27}$$

where $\tilde{s} > 0$ with $\tilde{s} \to 0$ and φ_i are given by (12) and (13); consequently, if we define

$$I_{min}(\eta) = \min\{(\eta - \varphi_0)\tilde{s}, (\varphi_1 - \eta)\tilde{s}\} \tag{28}$$

then

$$-\frac{1}{n}\ln E[C] \geq I_{min}(\eta). \tag{29}$$

Equivalently, we have the following upper bounds on the error probabilities and the average cost

$$P_0(\ln L_n > \eta) \leq \exp\{-n(\eta - \varphi_0)\tilde{s}\}; \quad P_1(\ln L_n \leq \eta) \leq \exp\{-n(\varphi_1 - \eta)\} \tag{30}$$

and

$$E[C] \leq \exp\{-nI_{min}(\eta)\} \tag{31}$$

for a large sample size n. As $n \to \infty$, these upper bounds tend to be tight in the sense that both the bounds and the error probabilities, and thus the average cost, converge to zero under the consistency condition $\varphi_0 \leq \eta \leq \varphi_1$. Then the optimal threshold is obtained by maximizing $I_{min}(\eta)$ as

$$\eta = \arg \max_{\varphi_0 \leq \eta \leq \varphi_1} I_{min}(\eta) = \arg \max_{\varphi_0 \leq \eta \leq \varphi_1} \min\{(\eta - \varphi_0)\tilde{s}, (\varphi_1 - \eta)\tilde{s}\}. \tag{32}$$

Since under the consistency condition $\varphi_0 \leq \eta \leq \varphi_1$ and for positive $\tilde{s}$, $(\eta - \varphi_0)\tilde{s}$ and $(\varphi_1 - \eta)\tilde{s}$ are increasing and decreasing functions of η, respectively, there is a single solution of the equation

$$(\eta - \varphi_0)\tilde{s} = (\varphi_1 - \eta)\tilde{s} \tag{33}$$

for η, and it is the optimal threshold characterized by (32). By solving (33) we obtain the optimal threshold given by $\eta^* = (\varphi_0 + \varphi_1)/2$; by employing this optimal threshold we have the final form of the lower bound on the asymptotic rate of the average cost given by

$$\begin{aligned} I_{min}(\eta^*) &= \frac{\varphi_1 - \varphi_0}{2}\tilde{s} \\ &= \frac{1}{2(1-\rho_0^2)}\left[\frac{(\mu_{1,1}-\mu_{1,0})^2}{\sigma_{1,0}^2} + \frac{(\mu_{2,1}-\mu_{2,0})^2}{\sigma_{2,0}^2} - \frac{2\rho_0(\mu_{1,1}-\mu_{1,0})(\mu_{2,1}-\mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right]\frac{\tilde{s}}{2} \\ &\quad + \frac{1}{2(1-\rho_1^2)}\left[\frac{(\mu_{1,1}-\mu_{1,0})^2}{\sigma_{1,1}^2} + \frac{(\mu_{2,1}-\mu_{2,0})^2}{\sigma_{2,1}^2} - \frac{2\rho_1(\mu_{1,1}-\mu_{1,0})(\mu_{2,1}-\mu_{2,0})}{\sigma_{1,1}\sigma_{2,1}}\right]\frac{\tilde{s}}{2}. \end{aligned} \tag{34}$$

Since $\tilde{s}/2$ is a positive constant, we may normalize the above $I_{min}(\eta^*)$ by dividing it by $\tilde{s}/2$ to obtain the design criteria for the optimal nonlinearities or quantizers given by

$$\begin{aligned} J_{min}(\varphi_0, \varphi_1) &= 2I_{min}(\eta^*)/\tilde{s} = \varphi_1 - \varphi_0 \\ &= \frac{1}{2(1-\rho_0^2)}\left[\frac{(\mu_{1,1}-\mu_{1,0})^2}{\sigma_{1,0}^2} + \frac{(\mu_{2,1}-\mu_{2,0})^2}{\sigma_{2,0}^2} - \frac{2\rho_0(\mu_{1,1}-\mu_{1,0})(\mu_{2,1}-\mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right] \\ &\quad + \frac{1}{2(1-\rho_1^2)}\left[\frac{(\mu_{1,1}-\mu_{1,0})^2}{\sigma_{1,1}^2} + \frac{(\mu_{2,1}-\mu_{2,0})^2}{\sigma_{2,1}^2} - \frac{2\rho_1(\mu_{1,1}-\mu_{1,0})(\mu_{2,1}-\mu_{2,0})}{\sigma_{1,1}\sigma_{2,1}}\right]. \end{aligned} \tag{35}$$

Thus the optimal nonlinearities or quantizers in Schemes 1-3 are characterized by the following maximization problem:

$$\begin{aligned} &\max_{g_1, g_2} J_{min}(\varphi_0, \varphi_1) \\ &\text{subject to } \mu_{k,1} \geq \mu_{k,0}; \quad k = 1, 2. \end{aligned} \tag{36}$$

In general, numerical optimization techniques, such as nonlinear programming, are required for solving such as maximization problem, and the two optimal nonlinearities or quantizers are coupled. In the cases treated in Sections III and IV, which are of practical interest, this maximization problem can be solved with the help of analytical methods.

2. Asymptotic Rate of Error Probabilities for Scheme 4

In Scheme 4, each sensor transmits its decision, a binary random variable $d_{n,k}$, $(k = 1, 2)$ to the fusion center. This random variable $d_{n,k}$ is not Gaussian distributed any more (not even asymptotically Gaussian as $n \to \infty$, and thus the design criterion derived above is not applicable. To derive a useful design criterion for the optimal nonlinearities, we introduce here the following asymptotic rates:

$$-\frac{1}{n}\ln P_0(d_{n,1}=1, d_{n,2}=1) = -\frac{1}{n}\ln P_0(T_{n,k} > \eta_1, T_{n,k} > \eta_2) \tag{37}$$

$$-\frac{1}{n}\ln P_1(d_{n,1}=0, d_{n,2}=0) = -\frac{1}{n}\ln P_1(T_{n,k} \leq \eta_1, T_{n,k} \leq \eta_2) \tag{38}$$

$$-\frac{1}{n}\ln P_i(d_{n,1}=1, d_{n,2}=0) = -\frac{1}{n}\ln P_i(T_{n,k} > \eta_1, T_{n,k} \leq \eta_2) \tag{39}$$

and

$$-\frac{1}{n}\ln P_i(d_{n,1}=0, d_{n,2}=1) = -\frac{1}{n}\ln P_i(T_{n,k} \leq \eta_1, T_{n,k} > \eta_2) \tag{40}$$

for $i = 0, 1$. In (37)-(40), $d_{n,k} = 0$, if sensor k decides in favor of H_0, and $d_{n,k} = 1$, if it decides in favor of H_1. The lemma that follows characterizes the closed-forms of the above asymptotic rates; it is taken from Theorem 5 of [8] and proved there with the help of the large deviation principle.

Lemma 2: Suppose

$$\frac{\eta_1 - \mu_{1,0}}{\sigma_{1,0}} \geq \rho_0 \frac{\eta_2 - \mu_{2,0}}{\sigma_{2,0}}, \quad \frac{\eta_2 - \mu_{2,0}}{\sigma_{2,0}} \geq \rho_0 \frac{\eta_1 - \mu_{1,0}}{\sigma_{1,0}} \tag{41}$$

and

$$\frac{\mu_{1,1} - \eta_1}{\sigma_{1,1}} \geq \rho_1 \frac{\mu_{2,2} - \eta_2}{\sigma_{2,1}}, \quad \frac{\mu_{2,2} - \eta_2}{\sigma_{2,1}} \geq \rho_1 \frac{\mu_{1,1} - \eta_1}{\sigma_{1,1}}. \tag{42}$$

Then, for jointly Gaussian distributed statistics $(T_{n,1}, T_{n,2})$, the asymptotic rates defined by (37)-(40) are given by

$$-\frac{1}{n}\ln P_0(d_{n,1}=0, d_{n,2}=1) \to \frac{(\eta_2 - \mu_{2,0})^2}{2\sigma_{2,0}^2}, \quad -\frac{1}{n}\ln P_1(d_{n,1}=0, d_{n,2}=1) \to \frac{(\mu_{1,1} - \eta_1)^2}{2\sigma_{1,1}^2} \tag{43}$$

$$-\frac{1}{n}\ln P_0(d_{n,1}=1,d_{n,2}=0)\to\frac{(\eta_1-\mu_{1,0})^2}{2\sigma_{1,0}^2},\quad -\frac{1}{n}\ln P_1(d_{n,1}=1,d_{n,2}=0)\to\frac{(\mu_{2,1}-\eta_2)^2}{2\sigma_{2,1}^2} \tag{44}$$

and

$$\begin{aligned}&-\frac{1}{n}\ln P_0(d_{n,1}=1,d_{n,1}=1)\\&\to\frac{1}{2(1-\rho_0^2)}\left[\frac{(\eta_1-\mu_{1,0})^2}{\sigma_{1,0}^2}+\frac{(\eta_2-\mu_{2,0})^2}{\sigma_{2,0}^2}-\frac{2\rho_0(\eta_1-\mu_{1,0})(\eta_2-\mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right]\end{aligned} \tag{45}$$

$$\begin{aligned}&-\frac{1}{n}\ln P_1(d_{n,1}=0,d_{n,1}=0)\\&\to\frac{1}{2(1-\rho_1^2)}\left[\frac{(\mu_{1,1}-\eta_1)^2}{\sigma_{1,1}^2}+\frac{(\mu_{2,1}-\eta_2)^2}{\sigma_{2,1}^2}-\frac{2\rho_1(\mu_{1,1}-\eta_1)(\mu_{2,1}-\eta_2)}{\sigma_{1,1}\sigma_{2,1}}\right]\end{aligned} \tag{46}$$

as $n\to\infty$. Furthermore,

$$-\frac{1}{n}\ln P_0(d_{n,1}=1,d_{n,1}=1)\geq-\frac{1}{n}\ln P_0(d_{n,1}=1,d_{n,1}=0) \tag{47}$$

$$-\frac{1}{n}\ln P_0(d_{n,1}=1,d_{n,1}=1)\geq-\frac{1}{n}\ln P_0(d_{n,1}=0,d_{n,1}=1) \tag{48}$$

$$-\frac{1}{n}\ln P_1(d_{n,1}=0,d_{n,1}=0)\geq-\frac{1}{n}\ln P_1(d_{n,1}=1,d_{n,1}=0) \tag{49}$$

$$-\frac{1}{n}\ln P_1(d_{n,1}=0,d_{n,1}=0)\geq-\frac{1}{n}\ln P_1(d_{n,1}=0,d_{n,1}=1). \tag{50}$$

In the following two sections, Lemma 2 is used to derive the design criterion for the optimal sensor nonlinearities and thresholds of Scheme 4. For notational convenience, we denote by $P_i(0,1)$, $P_i(1,0)$ $(i=0,1)$, $P_0(1,1)$, and $P_1(1,1)$ the probabilities $P_i(d_{n,1}=0,d_{n,2}=1)$, $P_i(d_{n,1}=1,d_{n,2}=0)$, $P_0(d_{n,1}=1,d_{n,2}=1)$, and $P_1(d_{n,1}=0,d_{n,2}=0)$, respectively.

III. DEPENDENCE ACROSS TIME

In this section, we consider the case of dependence across time, in which $\{X_l^{(1)}\}_{l=1}^n$ and $\{X_l^{(2)}\}_{l=1}^n$ are mutually independent. In this case, $\rho_1=\rho_0=0$ and the objective function $J_{min}(\varphi_0,\varphi_1)$ characterized in Subsection II.1 has the form

$$\begin{aligned}J_{min}(\varphi_0,\varphi_1)&=\left[\frac{1}{2\sigma_{1,0}^2}+\frac{1}{2\sigma_{1,1}^2}\right](\mu_{1,1}-\mu_{1,0})^2+\left[\frac{1}{2\sigma_{2,0}^2}+\frac{1}{2\sigma_{2,1}^2}\right](\mu_{2,1}-\mu_{2,0})^2\\&=J_1(\mu_{1,0},\mu_{1,1},\sigma_{1,0},\sigma_{1,1})+J_2(\mu_{2,0},\mu_{2,1},\sigma_{2,0},\sigma_{2,1})\end{aligned} \tag{51}$$

where J_k $(k=1,2)$ are defined by

$$J_k(\mu_{k,0},\mu_{k,1},\sigma_{k,0},\sigma_{k,1})=\left[\frac{1}{2\sigma_{k,0}^2}+\frac{1}{2\sigma_{k,1}^2}\right](\mu_{k,1}-\mu_{k,0})^2. \tag{52}$$

We notice that J_{min} given by (51) is the sum of J_1 and J_2, which are functionals of their own individual nonlinearities or quantizers of the two sensors. Therefore, the optimization problem of J_{min} characterized by (36) is decoupled into two optimization problems as

$$\begin{aligned} &\max_{g_k} J_k(\mu_{k,0}, \mu_{k,1}, \sigma_{k,0}, \sigma_{k,1}); \quad k = 1, 2 \\ &\text{subject to } \mu_{k,1} \geq \mu_{k,0} \end{aligned} \tag{53}$$

As discussed in Appendix B, this optimization problem can only be solved via numerical techniques, which are typically very computation-intensive and make the design procedure of the optimal quantizers (breakpoints and levels) extremely complex. To decrease the design complexity, we introduce here the following simplified form, for each J_k,

$$\bar{J}_k(\mu_{k,0}, \mu_{k,1}, \sigma_{k,0}, \sigma_{k,1}) = \frac{(\mu_{k,1} - \mu_{k,0})^2}{\sigma_{k,0}^2 + \sigma_{k,1}^2} \tag{54}$$

which has been used successfully in single-sensor discrimination (see [13]). According to the inequality

$$\frac{1}{2\sigma_{k,0}^2} + \frac{1}{2\sigma_{k,1}^2} = \frac{\sigma_{k,0}^2 + \sigma_{k,1}^2}{2\sigma_{k,0}^2\sigma_{k,1}^2} \geq \frac{\sigma_{k,0}^2 + \sigma_{k,1}^2}{(\sigma_{k,0}^2 + \sigma_{k,1}^2)^2/2} = \frac{2}{\sigma_{k,0}^2 + \sigma_{k,1}^2} \tag{55}$$

we obtain

$$J_k \geq 2\bar{J}_k; \quad k = 1, 2,$$

with the equality satisfied, if and only if $\sigma_{k,1}^2 = \sigma_{k,0}^2$. This implies that

$$\begin{aligned} E[C] &\leq \exp\{-nI_{min}\} = \exp\{-nJ_{min}\tilde{s}/2\} = \exp\{-n(J_1 + J_2)\tilde{s}/2\} \\ &\leq \exp\{-n(\bar{J}_1 + \bar{J}_2)\tilde{s}\} \end{aligned} \tag{56}$$

for all n; the first inequality becomes tight as $n \to \infty$ and the second as $\sigma_{k,1}^2 \to \sigma_{k,0}^2$. Thus the simplified measures $\bar{J}_k$ are also involved in an upper bound on the average cost and the upper bound characterized by them becomes tight as $\sigma_{k,1}^2 - \sigma_{k,0}^2 \to 0$.

We notice that $\bar{J}_k$ has the form of the generalized signal-to-noise ratio of [16]; for the case of detection of a weak signal in additive noise, $\sigma_{k,1}^2 = \sigma_{k,0}^2$ and $J_k = 2\bar{J}_k$ takes the form of the deflection criterion.

In the following, $\bar{J}_k$ is employed as the design criterion for the optimal nonlinearity or quantizer of Schemes 1-3 described in Section I, and this optimization problem is characterized by

$$\begin{aligned} &\max_{g_k} \bar{J}_k(\mu_{k,0}, \mu_{k,1}, \sigma_{k,0}, \sigma_{k,1}) \\ &\text{subject to } \mu_{k,1} \geq \mu_{k,0}; \quad k = 1, 2. \end{aligned} \tag{57}$$

1. Fusion of Unquantized Sensor Test Statistics

We first consider the scheme which has a configuration described in Figure 1. The sensor test statistics are fused without previous quantization. In this case, the maximization characterized by (57) and (54) was solved for single-sensor systems and continuous nonlinearities in [13]. Each optimal nonlinearity is obtained by solving the linear integral equation described by (see [13])

$$(f_{k,1}(x) - f_{k,0}(x))/(f_{k,1}(x) + f_{k,0}(x)) - \int \bar{K}_k(x,y) g_k(y) dy = g_k(x); \quad k = 1,2 \tag{58}$$

where

$$\begin{aligned} \bar{K}_k(x,y) &= \{2 \sum_{j=1}^{m_{k,1}} [f_{k,1}^{(j)}(x,y) - f_{k,1}(x) f_{k,1}(y)] + 2 \sum_{j=1}^{m_{k,0}} [f_{k,0}^{(j)}(x,y) - f_{k,0}(x) f_{k,0}(y)] \\ &\quad - f_{k,1}(x) f_{k,1}(y) - f_{k,0}(x) f_{k,0}(y)\} / \{f_{k,1}(x) + f_{k,0}(x)\} \end{aligned} \tag{59}$$

is the kernel of integration. For the case of m-dependent noise, the above kernel is well defined; for the ρ-mixing case we should take $m_{k,i} \to \infty$ but we also need to assume that the kernel in (59) exists, so as to be able to interchange summation and integration in (58).

2. Fusion of Suboptimally Quantized Sensor Observations

Here we consider Scheme 2, in which the nonlinearity used to quantize the sensor observations before transmitting them to the fusion center is obtained by discretizing the continuous nonlinearity g_k $(k = 1,2)$ of Scheme 1. In the discretized form of the integral equation given by (58), let the integration range be $(X_{min}^{(k)}, X_{max}^{(k)})$, which is assigned according to the support of the observation processes. Then

$$\frac{f_{k,1}(x_i^{(k)}) - f_{k,0}(x_i^{(k)})}{f_{k,1}(x_i^{(k)}) + f_{k,0}(x_i^{(k)})} - \sum_{j=0}^{M_k} \bar{K}_k(x_i^{(k)}, x_j^{(k)}) g_k(x_j^{(k)}) \Delta x_j^{(k)} = g_k(x_i^{(k)}); \quad k = 1,2 \tag{60}$$

where $x_i^{(k)}, i = 0,1,2,\cdots,M_k$ are the quantizer breakpoints with $x_0^{(k)} = X_{min}^{(k)}$ and $x_{M_k}^{(k)} = X_{max}^{(k)}$ and $\Delta x_j^{(k)}$ is the discrete approximation to dx (for example, $\Delta x_j^{(k)} = (X_{max}^{(k)} - X_{min}^{(k)})/M_k$), for sensor k $(k = 1,2)$.

Define the vectors

$$\underline{f_{k,i}} = \left[f_{k,i}(x_0^{(k)}),\ f_{k,i}(x_1^{(k)}), \cdots, f_{k,i}(x_{M_k}^{(k)})\right];\ \ i = 0,1 \tag{61}$$

and the matrix

$$\underline{\bar{G}_k} = [\bar{G}_{ij}^{(k)}] \tag{62}$$

where

$$\bar{G}_{ij}^{(k)} = (f_{k,1}(x_i^{(k)}) - f_{k,0}(x_i^{(k)}))[\bar{K}_k(x_i^{(k)}, x_j^{(k)})\Delta x_j^{(k)} + \delta(x_i^{(k)}, x_j^{(k)})] \tag{63}$$

with $\delta(x, y) = 1$, if $x = y$, and 0, otherwise. Then the equation (60) can be written as

$$(\underline{f_{k,1}} - \underline{f_{k,0}})^T = \underline{\bar{G}_k}\ \underline{g_k}^T;\ \ k = 1, 2 \tag{64}$$

where $\underline{g_k}$ is defined by

$$\underline{g_k} = \left[g_k(x_0^{(k)}), g_k(x_1^{(k)}), \cdots, g_k(x_{M_k}^{(k)})\right] \tag{65}$$

To solve (64) for $\underline{g_k}$, we assume that $x_i^{(k)}$ $(k = 1, 2;\ i = 0, 1, \cdots, M_k)$ are chosen for each sensor so that the matrix $\underline{\bar{G}_k}$ is nonsingular. Thus we obtain

$$\underline{g_k}^T = \underline{\bar{G}_k}^{-1}(\underline{f_{k,1}} - \underline{f_{k,0}})^T;\ \ k = 1, 2 \tag{66}$$

Then, for any observation x, we can characterize the quantizer of sensor k as follows:

$$\bar{g}_k(x) = \begin{cases} [g_k(x_0^{(k)}) + g_k(x_1^{(k)})]/2 & \text{if } x \le x_1^{(k)} \\ [g_k(x_i^{(k)}) + g_k(x_{i+1}^{(k)})]/2 & \text{if } x \in (x_i^{(k)}, x_{i+1}^{(k)}] \\ [g_k(x_{M_k-1}^{(k)}) + g_k(x_{M_k}^{(k)})]/2 & \text{if } x > x_{M_k-1}^{(k)} \end{cases} \tag{67}$$

for $i = 1, 2, \cdots, M_k - 2$.

Note that, in the above quantization scheme, the number of levels M_k and the breakpoints $x_i^{(k)}$, for $i = 0, 1, \cdots, M_k$, need not be the same for the different sensors. Moreover, the spacing of the breakpoints in the interval of the support of the noise process of sensor k need not be uniform. For the purpose of simplicity in analysis and implementation, we may set the number of quantization levels to be the same for all sensors and the breakpoints uniformly spaced over the interval of support. Although the resulting quantization scheme will not be optimal, its performance, as quantified from simulations, is acceptable for a reasonable number of quantization levels.

3. Fusion of Optimally Quantized Sensor Observations

In Scheme 2 of the previous subsection, the quantizer $\bar{g}_k$ of sensor k is obtained directly from the discrete-form of the optimal nonlinearity g_k of the Scheme 1 and thus is not an optimal quantizer. Since quantizers also function as nonlinearities [with their sum in (7) satisfying the central limit theorem] we can recompute the generalized signal-to-noise ratio $\bar{J}_k$ of (54) for sensor k by using quantizers and the test statistics of (7); then we can use it as the design criterion

for the optimal quantizer (breakpoints and levels) of each sensor. This type of scheme has been addressed in [14] and [15] for single-sensor detection of weak signals and a different performance measure (the efficacy). The technique of [15] can be extended from the single-sensor weak-signal case to the single-sensor discrimination case treated in [13] and then to the multi-sensor case treated below.

Let $Q_k(X_l^{(k)})$ $(l = 1, 2, \cdots, n)$ be the quantizer output, for sensor k $(k = 1, 2)$, when M_k quantization levels are employed. Denote by $\underline{t_k} = [t_{k,0}\ t_{k,1}\ \cdots\ t_{k,M_k}]$ its breakpoints and by $\underline{u_k} = [u_{k,1}\ u_{k,2}\ \cdots u_{k,M_k}]$ its levels. Then the performance measure $\bar{J}_k$ of sensor k $(k = 1, 2)$ in (54) can be formulated as a function of $\underline{u_k}$ and $\underline{t_k}$ by considering the quantizer Q_k as a special nonlinearity.

The variances $\tilde{\sigma}_{k,i}^2/n$ of $\frac{1}{n}\sum_{l=1}^{n} Q_k(X_l^{(k)})$ under $H_i^{(k)}$ have the form

$$\begin{aligned}\tilde{\sigma}_{k,i}^2 &= \sum_{l=1}^{M_k}(u_{k,l})^2 \int_{t_{k,l-1}}^{t_{k,l}} f_k(x)dx + 2\sum_{j=1}^{m_{k,i}}\sum_{r=1}^{M_k}\sum_{l=1}^{M_k} u_{k,r}u_{k,l}\int_{t_{k,r-1}}^{t_{k,r}}\int_{t_{k,l-1}}^{t_{k,l}} f_{k,i}^{(j)}(x,y)dxdy \\ &\quad -(2m_{k,i}+1)\left[\sum_{l=1}^{M_k} u_{k,l}\int_{t_{k,l-1}}^{t_{k,l}} f_k(x)dx\right]^2 \end{aligned} \tag{68}$$

while the corresponding means have the form

$$\tilde{\mu}_{k,i} = \int Q_k(x) f_{k,i}(x)dx = \sum_{l=1}^{M_k} u_{k,l}\int_{t_{k,l-1}}^{t_{k,l}} f_{k,i}(x)dx. \tag{69}$$

Thus the maximization of (57) will now be with respect to $\underline{u_k}$ and $\underline{t_k}$.

For $l = 1, 2, \cdots, M_k$, $k = 1, 2$, and under $H_i^{(k)}$ $(i = 0, 1)$, define the vector

$$\underline{\Delta f_{k,i}} = [\Delta f_1^{(k,i)}, \Delta f_2^{(k,i)}, \cdots, \Delta f_{M_k}^{(k,i)}] \tag{70}$$

where

$$\Delta f_l^{(k,i)} = \int_{t_{k,l-1}}^{t_{k,l}} f_{k,i}(x)dx \tag{71}$$

the matrixes

$$\underline{F_{k,i}} = diag\{\int_{t_{k,0}}^{t_{k,1}} f_{k,i}(x)dx, \cdots, \int_{t_{k,M_k-1}}^{t_{k,M_k}} f_{k,i}(x)dx\} \tag{72}$$

$$\underline{P_{k,i}} = [P_{r,l}^{(k,i)}] \tag{73}$$

and

$$\underline{R_{k,i}} = [R_{r,l}^{(k,i)}] \tag{74}$$

where, for $r, l = 1, 2, \cdots, M_k$,

$$P_{r,l}^{(k,i)} = 2 \sum_{j=1}^{m_{k,i}} \int_{t_{k,r-1}}^{t_{k,r}} \int_{t_{k,l-1}}^{t_{k,l}} f_{k,i}^{(j)}(x,y) dx dy \tag{75}$$

and

$$R_{r,l}^{(k,i)} = (2m_{k,i} + 1) \int_{t_{k,r-1}}^{t_{k,r}} f_{k,i}(x) dx \cdot \int_{t_{k,l-1}}^{t_{k,l}} f_{k,i}(x) dx. \tag{76}$$

Then $\bar{J}_k$ $(k = 1, 2)$ assumes the form

$$\bar{J}_k(Q_k) = \frac{[\underline{u_k}(\underline{\Delta f_{k,1}} - \underline{\Delta f_{k,0}})^T]^2}{\underline{u_k}[\sum_{i=0}^{1}(\underline{F_{k,i}} + \underline{P_{k,i}} - \underline{R_{k,i}})]\underline{u_k}^T}. \tag{77}$$

In this scheme, we assume that $t_{k,l}$ $(l = 0, 1, \cdots, M_k)$ of sensor k are chosen, such that the matrix $\sum_{i=0}^{1}(\underline{F_{k,i}} + \underline{P_{k,i}} - \underline{R_{k,i}})$ is positive definite. Then the optimal quantization levels of sensor k for fixed breakpoints are

$$\underline{u_k}^T = \left[\sum_{i=0}^{1}(\underline{F_{k,i}} + \underline{P_{k,i}} - \underline{R_{k,i}})\right]^{-1} (\underline{\Delta f_{k,1}} - \underline{\Delta f_{k,0}})^T. \tag{78}$$

Upon substitution for the optimal levels given by (78) into $\bar{J}_k$, we obtain

$$\bar{J}_k = (\underline{\Delta f_{k,1}} - \underline{\Delta f_{k,0}}) \left[\sum_{i=0}^{1}(\underline{F_{k,i}} + \underline{P_{k,i}} - \underline{R_{k,i}})\right]^{-1} (\underline{\Delta f_{k,1}} - \underline{\Delta f_{k,0}})^T \tag{79}$$

which is now a function of the breakpoints $\underline{t_k}$ only. By using numerical optimization techniques (such as the gradient method) for the objective function $\bar{J}_k$ given by (79), the optimal breakpoints can be obtained; then the optimal quantization levels are determined by (78).

4. Fusion of Binary (Hard) Sensor Decisions

In Scheme 4, the log-likelihood ratio test at the fusion center takes the form (see [2])

$$\begin{aligned} \ln L_n &= \ln \frac{P_1(d_{n,1}, d_{n,2})}{P_0(d_{n,1}, d_{n,2})} = \sum_{k=1}^{2} \ln \frac{p_1(d_{n,k})}{p_0(d_{n,k})} \\ &= n[w_1(d_{n,1}) + w_2(d_{n,2})] \gtrless_{H_0}^{H_1} n\eta = 0 \end{aligned} \tag{80}$$

where

$$w_k(d_{n,k}) = \begin{cases} \frac{1}{n} \ln \frac{1-\beta_k}{\alpha_k} = w_{k,1} & \text{if } d_{n,k} = 1 \\ -\frac{1}{n} \ln \frac{1-\alpha_k}{\beta_k} = -w_{k,0} & \text{if } d_{n,k} = 0 \end{cases} \tag{81}$$

with α_k and β_k being the probabilities of false alarm and miss for sensor k, respectively, and $n\eta = 0$ the threshold of the fusion center, which corresponds to equiprobable hypotheses H_1

and H_0 (prior probabilities $p = 1/2$). Thus the LRT fusion rule is characterized by what is a weighted sum of the decisions of the sensors (compare with [2] and [9]).

Since, in all cases of practical interest, $\alpha_k + \beta_k < 1$ under a large sample size n, we have that $w_{k,i} > 0$ ($k = 1, 2$, $i = 0, 1$) for the $w_{k,i}$'s defined by (81). Furthermore, if we denote by η_k ($k = 1, 2$) the threshold of sensor k and assume that it satisfies the consistency condition $\mu_{k,0} \leq \eta_k \leq \mu_{k,\theta}$, under which the error probability of each sensor tends to zero as $n \to \infty$ according to the ergodic mean theorem. Actually, we may use the large deviations principle of [17] (which we applied in [8] in the context of multi-sensor detection systems without a fusion center and in [9] for multi-sensor systems with a fusion center for the detection of weak signals) to obtain

$$\begin{aligned} w_{k,1} &= \frac{1}{n} \ln \frac{1-\beta_k}{\alpha_k} = \frac{1}{n}[\ln(1-\beta_k) - \ln \alpha_k] \\ &\to -\frac{1}{n} \ln \alpha_k \to \frac{(\eta_k - \mu_{k,0})^2}{2\sigma_{k,0}^2} \end{aligned} \tag{82}$$

and

$$\begin{aligned} w_{k,0} &= -\frac{1}{n} \ln \frac{1-\alpha_k}{\beta_k} = \frac{1}{n}[\ln(1-\alpha_k) - \ln \beta_k] \\ &\to -\frac{1}{n} \ln \beta_k \to \frac{(\mu_{k,1} - \eta_k)^2}{2\sigma_{k,1}^2} \end{aligned} \tag{83}$$

for $n \to \infty$. Thus the optimal decision rule in the fusion center is given by

$$\begin{aligned} &if \quad d_{n,1} = d_{n,2} = 0, \quad decide \quad H_0 \\ &if \quad d_{n,1} = 0,\ d_{n,2} = 1, \quad -w_{1,0} + w_{2,1} \underset{H_0}{\overset{H_1}{\gtrless}} 0 \\ &if \quad d_{n,1} = 1,\ d_{n,2} = 0, \quad w_{1,1} - w_{2,0} \underset{H_0}{\overset{H_1}{\gtrless}} 0 \\ &if \quad d_{n,1} = d_{n,2} = 1, \quad decide \quad H_1. \end{aligned} \tag{84}$$

This rule is similar to the one derived in Subsection III.5 of [9], except that the weights given by (82) - (83) are different from the corresponding weights of [9].

For the decision policy given by (84) and under the consistency condition for each sensor, the probabilities $P_0(0,1)$, $P_0(1,0)$, $P_0(1,1)$, $P_1(0,1)$, $P_1(1,0)$, $P_1(0,0)$ of the sensor decision pairs $(0,1)$, $(1,0)$, $(1,1)$, and $(0,0)$ converge to 0, as n increases, and so does the error probability of the fusion center P_e. Thus it is meaningful to consider the asymptotic exponential rate of the

error probability for large n, $-\frac{1}{n}\ln P_e$, as the design criterion for the optimal sensor nonlinearities g_k and thresholds η_k. The error probability P_e has the form

$$P_e = pP_0\{w_1(d_{n,1}) + w_2(d_{n,2}) > 0\} + (1-p)P_1\{w_1(d_{n,1}) + w_2(d_{n,2}) \leq 0\} \tag{85}$$

where p is the a priori probability of H_0 ($p = 1/2$ in our case). The optimal fusion test characterized by (84) results in the coupling of the two optimal thresholds of the sensors when they disagree (see [9]). As discussed in [9], to avoid the complex optimization problem that can be solved only via numerical techniques, we consider a simpler suboptimal approach which is shown via simulation results to provide satisfactory performance. According to this, we try to make $w_{1,1} + w_{2,1}$ and $-w_{1,0} - w_{2,0}$ respectively as large and as small as possible, since this will maximize their distance on opposite sides of the fusion center threshold $\eta = 0$ and thus minimize the fusion center error probability. Specifically, we determine the sensor thresholds by maximizing the minimum of the weights given by (82) - (82), that is,

$$\begin{aligned}
\eta_k &= \arg \max_{\mu_{k,0} \leq \eta_k \leq \mu_{k,1}} \min\{w_{k,1}, w_{k,0}\} \\
&= \arg \max_{\mu_{k,0} \leq \eta_k \leq \mu_{k,1}} \min\left\{-\frac{1}{n}\ln\alpha_k, -\frac{1}{n}\ln\beta_k\right\} \\
&= \arg \max_{\mu_{k,0} \leq \eta_k \leq \mu_{k,1}} \min\left\{\frac{(\eta_k - \mu_{k,0})^2}{2\sigma_{k,0}^2}, \frac{(\mu_{k,1} - \eta_k)^2}{2\sigma_{k,1}^2}\right\} \\
&= \frac{\sigma_{k,0}\mu_{k,1} + \sigma_{k,1}\mu_{k,0}}{\sigma_{k,0} + \sigma_{k,1}}.
\end{aligned} \tag{86}$$

Employing the above thresholds, we obtain

$$w_{k,1} = w_{k,0} = w_k = \frac{(\mu_{k,1} - \mu_{k,0})^2}{2(\sigma_{k,0} + \sigma_{k,1})^2}; \qquad k = 1, 2 \tag{87}$$

Next we evaluate the asymptotic rate $-\frac{1}{n}\ln P_e$ by applying Lemma 2. In the case treated in this section $\rho_0 = \rho_1 = 0$, and thus conditions (41) and (42) are satisfied, which enables the application of Lemma 2. We thus have the following expressions for the error probabilities α and β of the fusion center under the two hypotheses (see also [9])

$$\begin{aligned}
\alpha &= I(w_1 + w_2 > 0)P_0(1,1) + I(w_1 + w_2 > 0)P_0(0,1) \\
&\quad + I(w_1 - w_2 > 0)P_0(1,0) + I(w_1 - w_2 > 0)P_0(0,0) \\
&= P_0(1,1) + I(w_1 + w_2 > 0)P_0(0,1) + I(w_1 - w_2 > 0)P_0(1,0)
\end{aligned} \tag{88}$$

and

$$\begin{aligned}\beta &= I(w_1+w_2\le 0)P_1(1,1)+I(w_1+w_2\le 0)P_1(0,1)\\ &\quad +I(w_1-w_2\le 0)P_1(1,0)+I(w_1-w_2\le 0)P_1(0,0)\\ &= I(w_1+w_2\le 0)P_1(0,1)+I(w_1-w_2\le 0)P_1(1,0)+P_1(0,0) \qquad (89)\end{aligned}$$

where $I(A)$ for a set A denotes the indicator function of A. To minimize the above error probabilities we must consider the following different cases:

Case (i) : $w_1 > w_2$ or

$$\frac{(\mu_{1,1}-\mu_{1,0})^2}{(\sigma_{1,0}+\sigma_{1,1})^2} \ge \frac{(\mu_{2,1}-\mu_{2,0})^2}{(\sigma_{2,0}+\sigma_{2,1})^2}. \qquad (90)$$

Then $w_{1,1} - w_{2,0} > 0 = \eta$ and $-w_{1,0} + w_{2,1} < 0 = \eta$, which implies that

$$\alpha = P_0(1,1)+P_0(1,0) \quad \text{and} \quad \beta = P_1(0,1)+P_1(0,0) \qquad (91)$$

and the error probability of the fusion center has the form

$$P_e = p[P_0(1,1)+P_0(1,0)]+(1-p)[P_1(0,1)]+P_1(0,0)]. \qquad (92)$$

Intuitively the condition $w_1 > w_2$ corresponds to the decision of sensor 1 being more reliable than that of sensor 2, and thus, in case of disagreement between the two sensors, the fusion center decides according to the decision of sensor 1; this is reflected in the forms of the error probabilities derived above.

Let P_max be the maximum of the various error probability terms in P_e given by (92), that is,

$$P_{max} = \max\{P_0(1,1), P_0(1,0), P_1(0,1), P_1(0,0)\}. \qquad (93)$$

This P_{max} dominates the asymptotic rate of P_e. Using the results of Lemma 2 we obtain

$$\begin{aligned}-\frac{1}{n}\ln P_e \;&\to\; -\frac{1}{n}\ln P_{max} = \min\left\{\frac{(\eta_1-\mu_{1,0})^2}{2\sigma_{1,0}^2}, \frac{(\mu_{1,1}-\eta_1)^2}{2\sigma_{1,1}^2}\right\}\\ &= \frac{(\mu_{1,1}-\mu_{1,0})^2}{(\sigma_{1,0}+\sigma_{1,1})^2} \qquad (94)\end{aligned}$$

where

$$\eta_1 = \arg \max_{\mu_{1,0}\le\eta_1\le\mu_{1,1}} \min\left\{\frac{(\eta_1-\mu_{1,0})^2}{2\sigma_{1,0}^2}, \frac{(\mu_{1,1}-\eta_1)^2}{2\sigma_{1,1}^2}\right\}.$$

This is in total agreement with the results in (85) and (86).

The maximization of the performance measure resulting from the above asymptotic rate was discussed in [12]-[13] in the context of single-sensor discrimination and requires the solution of a nonlinear integral equation with respect to the nonlinearity g_1. This nonlinear integral equation was shown in [12] to have a unique solution and can be efficiently solved as indicated in [13]. Still it is difficult to robustify to uncertainty in the statistics of the observations. An alternative approach is to use the inequality $(x+y)^2 \leq 2(x^2+y^2)$ (which becomes an equality for $x=y$) and obtain the following lower bound on the asymptotic rate:

$$\frac{(\mu_{1,1}-\mu_{1,0})^2}{(\sigma_{1,0}+\sigma_{1,1})^2} \geq \frac{(\mu_{1,1}-\mu_{1,0})^2}{2(\sigma_{1,0}^2+\sigma_{1,1}^2)}. \tag{95}$$

This lower bound becomes tight as $\sigma_{1,1}^2 \to \sigma_{1,0}^2$. The maximization of this lower bound with respect to g_1 has been addressed in Scheme 1 of Subsection III.1 and results in the solution of a linear integral equation. When $w_1 > w_2$, we may use the above lower bound in the design of the optimal nonlinearity g_1. Since we would like the inequality $w_1 > w_2$ to hold for our choice of g_1 as the nonlinearity maximizing (95), it suffices that the desirable inequality ($w_1 > w_2$) holds for the g_2 that maximizes $(\mu_{2,1}-\mu_{2,0})^2/(\sigma_{2,0}+\sigma_{2,1})^2$ (i.e., the weight w_2); this g_2 can also be obtained from the analysis of Subsection III.1.

Case (ii) : $w_1 \leq w_2$ or

$$\frac{(\mu_{1,1}-\mu_{1,0})^2}{(\sigma_{1,0}+\sigma_{1,1})^2} \leq \frac{(\mu_{2,1}-\mu_{2,0})^2}{(\sigma_{2,0}+\sigma_{2,1})^2}. \tag{96}$$

Then $w_1 - w_2 \leq 0$ and $-w_1 + w_2 \geq 0$. This case is dual to case (i). The formulation and necessary steps are similar to those of case (i); we only need exchange the roles of sensors 1 and 2 to obtain the appropriate optimization conditions for sensor nonlinearities and thresholds. Therefore, we do not repeat the details here.

In summary, for Scheme 4 and under equiprobable hypotheses the optimal thresholds and nonlinearities of sensors are determined from (86) and (58)-(59), respectively.

IV. DEPENDENCE ACROSS TIME AND SENSORS

The case of interest here is that of identical sensor univariate and bivariate densities, i.e., $f_{1,i}(x) = f_{2,i}(x) = f_i(x)$ and $f_{1,i}^{(j)}(x,y) = f_{2,i}^{(j)}(x,y) = f_i^{(j)}(x,y)$, for $i = 0,1$ and $j = 1,2,\cdots,m_i$, where we assume that $m_{1,i} = m_{2,i} = m_i$ for the range of dependence of the sensor observations. In addition, let $\tilde{f}_i^{(j)}(x,y) = \tilde{f}_i(X_1^{(1)} = x, X_{j+1}^{(2)} = y)$, for $j = 0,1,2,\cdots,\tilde{m}_i$, denote the joint densities of the two observation processes, where $\tilde{m}_i$ is the range of dependence of the observations across the two sensors under hypothesis H_i ($i = 0,1$). It is easy to verify that the choice

$g_1 = g_2 = g$ for the sensor nonlinearities satisfies the necessary conditions for the maximization problem given by (36). Although we cannot argue that this is the optimal solution, it is a reasonable choice, as it simplifies the implementation and provides a satisfactory performance (as we establish via simulation). The optimal solution is very difficult to obtain via analytical techniques.

Thus $\mu_{1,i} = \mu_{2,i} = \mu_i$ and $\sigma_{1,i} = \sigma_{2,i} = \sigma_i$, under $H_i^{(k)}$ $(i = 0, 1)$, and J_{min} takes the form

$$J_{min}(\varphi_0, \varphi_1) = (\mu_1 - \mu_0)^2 \left[\frac{1}{(1+\rho_0)\sigma_0^2} + \frac{1}{(1+\rho_1)\sigma_1^2} \right]. \tag{97}$$

Similar to the maximization problem characterized by (53), the maximization of the above J_{min} under $\mu_1 \gtrless \mu_0$ requires numerical techniques and is not amenable to further analysis. Hence, we turn to a lower bound on J_{min} similar to the one used in Section III. In particular, we apply a modified version of the inequality in (55) to deduce the following lower bound on J_{min}

$$J_{min} \geq 4\bar{J}$$

where

$$\bar{J}(\mu_i, \sigma_i, \rho_i; i = 0, 1) = \frac{(\mu_1 - \mu_0)^2}{(1+\rho_0)\sigma_0^2 + (1+\rho_1)\sigma_1^2}. \tag{98}$$

with the equality satisfied, if and only if $\sigma_1^2 = \sigma_0^2$ and $\rho_1 = \rho_0$. This implies that

$$\begin{aligned} E[C] &\leq \exp\{-nI_{min}\} = \exp\{-nJ_{min}\tilde{s}/2\} \\ &\leq \exp\{-n(2\bar{J})\tilde{s}\} \end{aligned}$$

for all n; the first inequality becomes tight as $n \to \infty$ and the second as $\sigma_1^2 \to \sigma_0^2$ and $\rho_1 \to \rho_0$. Then the optimal g is obtained by solving the maximization problem described by

$$\begin{aligned} &\max_g \bar{J}(\mu_i, \sigma_i, \rho_i; i = 0, 1) \\ &\text{subject to } \mu_1 \geq \mu_0. \end{aligned} \tag{99}$$

This maximization problem is solved in Appendix C, and the sufficient and the necessary conditions for the optimal g are derived there.

We notice that both (97) and (98) define new expressions for the signal-to-noise ratio of two-sensor discrimination systems. These performance measures are reminiscent of (but distinctly different from) the corresponding efficacy-type measures derived in [9] for the two-sensor

detection of weak signals. For the rest of this section, we use $\bar{J}$ as the design criterion for the nonlinearity or the quantizer in the various quantization and fusion schemes except for Scheme 4.

1. Fusion of Unquantized Sensor Test Statistics

From the maximization of $\bar{J}_{min}$ performed in Appendix C we obtain that the optimal nonlinearity satisfies

$$(f_1(x) - f_0(x))/(f_1(x) + f_0(x)) - \int \bar{K}_c(x,y)g(y)dy = g(x) \tag{100}$$

where the integration kernel takes the form

$$\begin{aligned}\bar{K}_c(x,y) \;=\; & \{2\sum_{j=1}^{m_1}[f_1^{(j)}(x,y) - f_1(x)f_1(y)] + 2\sum_{j=1}^{\tilde{m}_1}[\tilde{f}_1^{(j)}(x,y) - f_1(x)f_1(y)] \\ & +2\sum_{j=1}^{m_0}[f_0^{(j)}(x,y) - f_0(x)f_0(y)] + 2\sum_{j=1}^{\tilde{m}_0}[\tilde{f}_0^{(j)}(x,y) - f_0(x)f_0(y)] \\ & +\tilde{f}_1^{(0)}(x,y) - 2f_1(x)f_1(y) + \tilde{f}_0^{(0)}(x,y) - 2f_0(x)f_0(y)\}/\{f_1(x) + f_0(x)\}. \end{aligned} \tag{101}$$

This is similar to the derivation of g_k for Scheme 1 in Subsection III.1.

2. Fusion of Suboptimally Quantized Sensor Observations

Similar to Scheme 2 in Subsection III.2, we now obtain the discretized form of (100) as

$$\frac{f_1(x_i) - f_0(x_i)}{f_1(x_i) + f_0(x_i)} - \sum_{j=0}^{M} \bar{K}_c(x_i,x_j)g(x_j)\Delta x_j = g(x_i) \tag{102}$$

where x_i, for $i = 0, 1, 2, \cdots, M$ with $M = M_1 = M_2$, are the breakpoints. We define the vectors

$$\underline{f_i} = [f_i(x_0), f_i(x_1), \cdots, f_i(x_M)]; \;\; i = 0, 1 \tag{103}$$

and the matrix

$$\bar{\underline{G}} = [\bar{G}_{ij}] \quad . \tag{104}$$

where

$$\bar{G}_{ij} = (f_1(x_i) - f_0(x_i)[\bar{K}(x_i,x_j)\Delta x_j + \delta(x_i,x_j)]. \tag{105}$$

Then (102) can be written as

$$(\underline{f_1} - \underline{f_0})^T = \bar{\underline{G}}\,\underline{g}^T \tag{106}$$

where $\underline{g}$ is defined in the same way as $\underline{g_k}$ in Subsection III.2. We again assume that x_i $(i = 0, 1, \cdots, M)$ are chosen such that the matrix $\underline{\bar{G}}$ is nonsingular. Thus we have

$$\underline{g}^T = \underline{\bar{G}}^{-1}(\underline{f_1} - \underline{f_0})^T \tag{107}$$

and the quantizer for this scheme is characterized by

$$\bar{g}(x) = \begin{cases} [g(x_0) + g_k(x_1)]/2 & \text{if } x \leq x_1 \\ [g(x_i) + g(x_{i+1})]/2 & \text{if } x \in (x_i, x_{i+1}] \\ [g(x_{M-1}) + g(x_M)]/2 & \text{if } x > x_{M-1} \end{cases} \tag{108}$$

for $i = 1, 2, \cdots, M-2$.

3. Fusion of Optimally Quantized Sensor Observations

Similar to Scheme 3 of subsection III.3 but using the same structure for both quantizers, in other words, $Q_1 = Q_2 = Q$, $\underline{t_1} = \underline{t_2} = \underline{t}$ and $\underline{u_1} = \underline{u_2} = \underline{u}$, we obtain under $H_i^{(k)}$ $(i = 0, 1)$

$$\begin{aligned} \tilde{\rho}_i \tilde{\sigma}_i^2 &= E_i[Q(X_1^{(1)})Q(X_1^{(2)})] - (E_i[Q(X_1^{(1)})])^2 + 2\sum_{j=1}^{\tilde{m}_i}\{E_i[Q(X_1^{(1)})Q(X_{j+1}^{(2)})] - (E_i[Q(X_1^{(1)})])^2\} \\ &= \sum_{r=1}^{M}\sum_{l=1}^{M} u_r u_l \int_{t_{r-1}}^{t_r}\int_{t_{l-1}}^{t_l} \tilde{f}_i^{(0)}(x,y)dxdy + 2\sum_{j=1}^{\tilde{m}_i}\{\sum_{r=1}^{M}\sum_{l=1}^{M} u_r u_l \cdot \int_{t_{r-1}}^{t_r}\int_{t_{l-1}}^{t_l} \tilde{f}_i^{(j)}(x,y)dxdy\} \\ &\quad -(2\tilde{m}_i + 1)[\sum_{l=1}^{M} u_l \int_{t_{l-1}}^{t_l} f_i(x)dx]^2 \end{aligned} \tag{109}$$

where $\tilde{\rho}_i$ is the correlation coefficient $\tilde{T}_{n,1}$ and $\tilde{T}_{n,2}$. In addition,

$$\tilde{\mu}_i = \int Q(x) f_i(x)dx = \sum_{l=1}^{M} u_l \int_{t_{l-1}}^{t_l} f_i(x)dx. \tag{110}$$

Let the vectors $\underline{\Delta f_{k,i}}$ and the matrixes $\underline{F_{k,i}}$, $\underline{P_{k,i}}$ and $\underline{R_{k,i}}$ be defined in the same way as in Scheme 3 of Subsection III.3. Then under symmetric densities and identical quantizers, $\underline{\Delta f_{1,i}} = \underline{\Delta f_{2,i}} = \underline{\Delta f_i}$, $\underline{F_{1,i}} = \underline{F_{2,i}} = \underline{F_i}$, $\underline{P_{1,i}} = \underline{P_{2,i}} = \underline{P_i}$ and $\underline{R_{1,i}} = \underline{R_{2,i}} = \underline{R_i}$ $(i = 0, 1)$. Moreover, under $H_i^{(k)}$ $(i = 0, 1)$, define the matrixes $\underline{P_{c,i}} = [P_{c,r,l}^{(i)}]$ and $\underline{R_{c,i}} = [R_{c,r,l}^{(i)}]$ where

$$P_{c,r,l}^{(i)} = \int_{t_{r-1}}^{t_r}\int_{t_{l-1}}^{t_l} \tilde{f}_i^{(0)}(x,y)dxdy + 2\sum_{j=1}^{\tilde{m}_i}\int_{t_{r-1}}^{t_r}\int_{t_{l-1}}^{t_l} \tilde{f}_i^{(j)}(x,y)dxdy. \tag{111}$$

and

$$R_{c,r,l}^{(i)} = (2\tilde{m}_i + 1) \cdot \int_{t_{r-1}}^{t_r} f_i(x)dx \cdot \int_{t_{l-1}}^{t_l} f_i(x)dx \tag{112}$$

Then the objective function $\bar{J}_{min}$ of quantizer Q is given by

$$\bar{J}_{min}(Q) = \frac{[\underline{u}(\underline{\Delta f_1} - \underline{\Delta f_0})^T]^2}{\underline{u}[\sum_{i=0}^{1}(\underline{F_i} + \underline{P_i} + \underline{P_{c,i}} - \underline{R_i} - \underline{R_{c,i}})]\underline{u}^T}. \tag{113}$$

We also assume that $\underline{t}$ are such that the matrix $\sum_{i=0}^{1}(\underline{F_i} + \underline{P_i} + \underline{P_{c,i}} - \underline{R_i} - \underline{R_{c,i}})$ is positive definite. Then the optimal quantization levels of sensor k are given by

$$\underline{u}^T = \left[\sum_{i=0}^{1}(\underline{F_i} + \underline{P_i} + \underline{P_{c,i}} - \underline{R_i} - \underline{R_{c,i}})\right]^{-1} (\underline{\Delta f_1} - \underline{\Delta f_0})^T \tag{114}$$

for fixed breakpoints. Upon substitution for the above optimal levels into $\bar{J}_{min}$, we obtain $\bar{J}_{min}$ as a function of the breakpoints $\underline{t}$ as follows:

$$\bar{J}_{min} = (\underline{\Delta f_1} - \underline{\Delta f_0}) \left[\sum_{i=0}^{1}(\underline{F_i} + \underline{P_i} + \underline{P_{c,i}} - \underline{R_i} - \underline{R_{c,i}})\right]^{-1} (\underline{\Delta f_1} - \underline{\Delta f_0})^T. \tag{115}$$

The maximization of this expression with respect to the breakpoints can be accomplished via numerical optimization techniques.

4. Fusion of Binary (Hard) Sensor Decisions

Unlike Scheme 4 of Subsection III.4 and due to the correlation between the sensor decisions $d_{n,1}$ and $d_{n,2}$, the log-likelihood ratio function of the fusion center does not assume an additive form. The asymptotic log-likelihood ratio is given by

$$\frac{1}{n}\ln L_n = \frac{1}{n}\ln P_1(d_{n,1}, d_{n,2}) - \frac{1}{n}\ln P_0(d_{n,1}, d_{n,2}) \tag{116}$$

for any sample size n, and the corresponding log-likelihood ratio test with the threshold $\eta = 0$ (for equiprobable hypotheses) at the fusion center is given by

$$\frac{1}{n}\ln P_1(d_{n,1}, d_{n,2}) - \frac{1}{n}\ln P_0(d_{n,1}, d_{n,2}) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{117}$$

which involves the asymptotic rates of the probabilities $P_i(0,1)$, $P_i(1,0)$, $P_0(1,1)$, and $P_1(0,0)$ of the sensor decisions.

Suppose that the conditions (41) and (42) are satisfied. Then we can apply the results of Lemma 2 to our case and obtain that, for a large sample size n, the optimal fusion rule is described by

$$if\ d_{n,1} = 0, d_{n,2} = 1; \qquad \frac{1}{n}\ln P_1(0,1) - \frac{1}{n}\ln P_0(0,1)$$

$$\rightarrow \quad -\frac{(\mu_{1,1}-\eta_1)^2}{2\sigma_{1,1}^2}+\frac{(\eta_2-\mu_{2,0})^2}{2\sigma_{2,0}^2} \underset{H_0}{\overset{H_1}{\lessgtr}} 0$$

$$if\ d_{n,1}=1, d_{n,2}=0; \qquad \frac{1}{n}\ln P_1(1,0)-\frac{1}{n}\ln P_0(1,0)$$

$$\rightarrow \quad -\frac{(\mu_{2,1}-\eta_2)^2}{2\sigma_{2,1}^2}+\frac{(\eta_1-\mu_{1,0})^2}{2\sigma_{1,0}^2} \underset{H_0}{\overset{H_1}{\lessgtr}} 0 \tag{118}$$

$$if\ d_{n,1}=0, d_{n,2}=0; \qquad \frac{1}{n}\ln P_1(0,0)-\frac{1}{n}\ln P_0(0,0)$$

$$\rightarrow \quad -\frac{1}{2(1-\rho_1^2)}\left[\frac{(\mu_{1,1}-\eta_1)^2}{\sigma_{1,1}^2}+\frac{(\mu_{2,1}-\eta_2)^2}{\sigma_{2,1}^2}-\frac{2\rho_1(\mu_{1,1}-\eta_1)(\mu_{2,1}-\eta_2)}{\sigma_{1,1}\sigma_{2,1}}\right] \underset{H_0}{\overset{H_1}{\lessgtr}} 0$$

$$if\ d_{n,1}=1, d_{n,2}=1; \qquad \frac{1}{n}\ln P_1(1,1)-\frac{1}{n}\ln P_0(1,1)$$

$$\rightarrow \quad \frac{1}{2(1-\rho_0^2)}\left[\frac{(\eta_1-\mu_{1,0})^2}{\sigma_{1,0}^2}+\frac{(\eta_2-\mu_{2,0})^2}{\sigma_{2,0}^2}-\frac{2\rho_0(\eta_1-\mu_{1,0})(\eta_2-\mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right] \underset{H_0}{\overset{H_1}{\lessgtr}} 0.$$

where we use the fact that $P_1(1,1) \rightarrow 1$ and $P_0(0,0) \rightarrow 1$, as $n \rightarrow \infty$, under the consistency conditions $\mu_{k,0} \leq \eta_k \leq \mu_{k,\theta}$ for the sensor thresholds. Furthermore, notice that

$$\frac{1}{n}\ln P_1(0,0)-\frac{1}{n}\ln P_0(0,0) \leq 0$$

and

$$\frac{1}{n}\ln P_1(1,1)-\frac{1}{n}\ln P_0(1,1) > 0$$

in the last two subtests of (118), since $x^2+y^2-2\rho xy=(x-\rho y)^2+(1-\rho^2)y^2=(y-\rho x)^2+(1-\rho^2)x^2 \geq 0$, for all $\rho \in [-1,1]$. Therefore, the sensor decision pairs $(0,0)$ or $(1,1)$ always imply that the fusion center favors H_0 or H_1, respectively, in its decisions. To pursue the optimal sensor nonlinearities and associated thresholds for the general case of asymmetric (unequal) sensor observation densities is a very difficult task. Instead, as in the previous subsections, we restrict our attention to the case of symmetric univariate and bivariate densities.

In the symmetric case, we can easily check that the conditions (41) and (42) are satisfied, since $\rho_i \leq 1$ $(i=0,1)$, and the factors multiplying it in these conditions are nonnegative, because of the consistency conditions. Using the results of Lemma 2 we obtain the fusion rule

$$\begin{aligned}
&if\ d_{n,1}=0, d_{n,2}=1;\ -\frac{(\mu_1-\tilde{\eta})^2}{2\sigma_1^2}+\frac{(\tilde{\eta}-\mu_0)^2}{2\sigma_0^2} \underset{H_0}{\overset{H_1}{\lessgtr}} 0\\
&if\ d_{n,1}=1, d_{n,2}=0;\ -\frac{(\mu_1-\tilde{\eta})^2}{2\sigma_1^2}+\frac{(\tilde{\eta}-\mu_0)^2}{2\sigma_0^2} \underset{H_0}{\overset{H_1}{\lessgtr}} 0 \qquad (119)\\
&if\ d_{n,1}=0, d_{n,2}=0;\ -\frac{(\mu_1-\tilde{\eta})^2}{(1+\rho_1)\sigma_1^2} \underset{H_0}{\overset{H_1}{\lessgtr}} 0\\
&if\ d_{n,1}=1, d_{n,2}=1;\ \frac{(\tilde{\eta}-\mu_0)^2}{(1+\rho_0)\sigma_0^2} \underset{H_0}{\overset{H_1}{\lessgtr}} 0
\end{aligned}$$

where $\tilde{\eta} = \eta_1 = \eta_2$ is the threshold used by the two sensors. This threshold is determined in the manner described by (86) in Subsection III.4, that is by maximizing $\min\{(\tilde{\eta} - \mu_0)^2/[(1 + \rho_0)\sigma_0^2], (\mu_1 - \tilde{\eta})^2/[(1 + \rho_1)\sigma_1^2]\}$ [the minimum of the weights for sensor decision pairs $(0,0)$ and $(1,1)$], except that the final result is now the following:

$$\tilde{\eta} = \frac{(1 + \rho_0)\sigma_0\mu_1 + (1 + \rho_1)\sigma_1\mu_0}{(1 + \rho_0)\sigma_0 + (1 + \rho_1)\sigma_1}. \tag{120}$$

With this threshold, the fusion center subtests for the various sensor decision pairs become

$$\begin{aligned}
&if\ d_{n,1} = 0, d_{n,2} = 1;\ \frac{(\mu_1 - \mu_0)^2(2 + \rho_1 + \rho_0)}{[(1 + \rho_1)\sigma_1 + (1 + \rho_0)\sigma_0]^2}(-\rho_1 + \rho_0) \underset{H_0}{\overset{H_1}{\gtrless}} 0,\ \text{i.e.},\ -\rho_1 + \rho_0 \underset{H_0}{\overset{H_1}{\gtrless}} 0 \\
&if\ d_{n,1} = 1, d_{n,2} = 0;\ \frac{(\mu_1 - \mu_0)^2(2 + \rho_1 + \rho_0)}{[(1 + \rho_1)\sigma_1 + (1 + \rho_0)\sigma_0]^2}(-\rho_1 + \rho_0) \underset{H_0}{\overset{H_1}{\gtrless}} 0,\ \text{i.e.},\ -\rho_1 + \rho_0 \underset{H_0}{\overset{H_1}{\gtrless}} 0 \qquad (121) \\
&if\ d_{n,1} = 0, d_{n,2} = 0;\ -\frac{(\mu_1 - \mu_0)^2}{[(1 + \rho_1)\sigma_1 + (1 + \rho_0)\sigma_0]^2} \underset{H_0}{\overset{H_1}{\gtrless}} 0,\ \text{i.e., decide } H_0 \\
&if\ d_{n,1} = 1, d_{n,2} = 1;\ \frac{(\mu_1 - \mu_0)^2}{[(1 + \rho_1)\sigma_1 + (1 + \rho_0)\sigma_0]^2} \underset{H_0}{\overset{H_1}{\gtrless}} 0,\ \text{i.e., decide } H_1.
\end{aligned}$$

The error probability of the fusion center takes the form

$$\begin{aligned}
P_e &= p\{P_0(1,1) + I(-\rho_1 + \rho_0 > 0)P_0(0,1) + I(-\rho_1 + \rho_0 > 0)P_0(1,0)\} \\
&\quad + (1 - p)\{P_1(0,0) + I(-\rho_1 + \rho_0 \le 0)P_1(0,1) + I(-\rho_1 + \rho_0 \le 0)P_1(1,0)\}
\end{aligned} \tag{122}$$

where $I(A)$ is the indicator function of A. To minimize this error probability, we consider the cases (i) $\rho_1 > \rho_0$ and (ii) $\rho_1 \le \rho_0$. Next we only discuss case (i).

In the case $\rho_1 > \rho_0$, the error probability of the fusion center becomes

$$P_e = pP_0(1,1) + (1 - p)[P_1(0,0) + P_1(0,1) + P_1(1,0)] \tag{123}$$

whose asymptotic rate is derived by using the results of Lemma 2 as

$$\begin{aligned}
-\frac{1}{n}\ln P_e &\to \min\left\{\frac{(\tilde{\eta} - \mu_0)^2}{2\sigma_0^2}, \frac{(\mu_1 - \tilde{\eta})^2}{2\sigma_1^2}, \frac{(\tilde{\eta} - \mu_0)^2}{(1 + \rho_0)\sigma_0^2}, \frac{(\mu_1 - \tilde{\eta})^2}{(1 + \rho_1)\sigma_1^2}\right\} \\
&= \min\left\{\frac{(\tilde{\eta} - \mu_0)^2}{2\sigma_0^2}, \frac{(\mu_1 - \tilde{\eta})^2}{2\sigma_1^2}\right\} = \frac{(\mu_1 - \mu_0)^2}{(\sigma_0 + \sigma_1)^2}
\end{aligned} \tag{124}$$

as $n \to \infty$. Notice that a different sensor threshold, namely

$$\tilde{\eta} = (\sigma_0\mu_1 + \sigma_1\mu_0)/(\sigma_0 + \sigma_1) \tag{125}$$

maximizes the above asymptotic rate. If we use this threshold in the fusion rule of (119), we obtain a fusion rule different from that of (121). Actually, since the fusion center error

probability is the ultimate performance criterion, the above choice of threshold in (125) and the fusion rule of (119) with this threshold in place should be preferred over any other design. We presented here the alternative design based on the maximization of the weights in the fusion rule of (119) for the sake of completeness. The fact is that no matter which specific inequalities are satisfied in the various subtests of the fusion rule of (119) (because of the choice of $\tilde{\eta}$), the error probability of the fusion center will have the asymptotic rate given by (124).

The asymptotic rate of (124) is lower bounded by $(\mu_1 - \mu_0)^2/2(\sigma_0^2 + \sigma_1^2)$; this lower bound becomes tight as $\sigma_1^2 \to \sigma_0^2$. This performance measure is preferable, since its maximization results in a linear integral equation, as already discussed at the beginning of Section III. Therefore, the sensor nonlinearity g is determined as the solution of the optimization problem described in (57).

V. PERFORMANCE EVALUATION OF TWO-SENSOR DISCRIMINATION SCHEMES

In this section we compare via simulation the performance of the various quantization and fusion schemes which take into account the dependence in the sensor observations and employ the optimal nonlinearities or quantizers to the ones which ignore the dependence. Moreover, we simulate and quantify the relative performance of the four quantization and fusion schemes introduced and analyzed in the previous sections. The sensor observation processes are characterized by stationary lognormal (under H_1) and Rayleigh (under H_0) univariate and bivariate (second-order joint) densities. Actually, a ρ-mixing dependence model is adopted for the sensor observations. These distributions under two hypotheses are motivated by practical problems in naval target discrimination, in which H_1 models a target (e.g., a ship) and H_0 a decoy (e.g., a chaff cloud). It is assumed that the first- and second-order moments of the sensor observations under H_0 are equal to the corresponding ones under H_1; which makes the discrimination problem particularly difficult. The two hypotheses are assumed to be equiprobable, i.e., $p = .5$. A sample size of $n = 600$ is used in all examples to yield sufficiently small error probabilities. 1000 simulations are run to generate the numerical results in this section; this is sufficient for providing reliable results, as long as the error probabilities are no smaller than .005, which is the case here. In our simulations the magnitude of the sensor observations is restricted to be in the interval $[X_{min}, X_{max}]$, this is imposed from practical considerations involving the hardware of the radar discrimination systems. In this way, very small and in particular very large samples (in the tails of the pdfs) are discarded and good quality simulation data are fed to our

quantization and fusion schemes. The appropriate values of X_{min} and X_{max} are derived for the specific sensor observation pdfs, by requiring that $P_{k,i}\{X \leq X_{min}\} = P_{k,i}\{X \geq X_{max}\} \leq \epsilon$ for $i = 0, 1$, $k = 1, 2$, and $\epsilon = 10^{-6}$. For the lognormal and Rayleigh univariate and bivariate pdfs of interest in this section the range of observations turns out to be $(0.02, 16.6)$. Finally, in all examples involving quantizers, we use $M_1 = M_2 = 8$, i.e., eight-level quantizers; the number of quantization levels was kept small to reduce the complexity of the computation of the optimal breakpoints and quantization levels.

The lognormal processes under H_1, $X_{1,i}^{(k)}$ for $i = 1, 2, \cdots, n$ and $k = 1, 2$ are obtained from the nonlinear transformations $X_{1,i}^{(k)} = \exp[\sigma_{G,1}^{(k)} N_{1,i}^{(k)} + \mu_{G,1}^{(k)}]$, in which the Gaussian processes $N_{1,i}^{(k)}$ are generated from the recursion formula

$$\begin{aligned} N_{1,1}^{(k)} &= V_{1,1}^{(k)} \\ N_{1,i}^{(k)} &= \rho_{G,1}^{(k)} N_{1,i-1}^{(k)} + \sqrt{1 - [\rho_{G,1}^{(k)}]^2} V_{1,i}^{(k)};\ i > 1 \end{aligned}$$

where $-1 \leq \rho_{G,1}^{(k)} \leq 1$ are correlation coefficients and $V_{1,i}^{(k)}$ $(i = 1, 2, \cdots, n)$, for sensor k $(k = 1, 2)$, is a sequence of i.i.d. standard Gaussian $(\mathcal{N}(0, 1))$ random variables. In the case of dependence across time only, $V_{1,i}^{(1)}$ and $V_{1,i}^{(2)}$ are generated independently; in the case of dependence across time and sensors, they are generated in a coupled manner as

$$V_{1,i}^{(2)} = \rho_{G,1}^{(c)} V_{1,i}^{(1)} + \sqrt{1 - [\rho_{G,1}^{(c)}]^2} W_{1,i}$$

where $W_{1,i}$ is another i.i.d. standard Gaussian $(\mathcal{N}(0, 1))$ process and $-1 \leq \rho_{G,1}^{(c)} \leq 1$ a correlation parameter; $V_{1,i}^{(1)}$ and $W_{1,i}$ are mutually independent.

The Rayleigh processes under H_0, $X_{0,i}^{(k)}$ are obtained from two independent Gaussian processes $Y_{0,i}^{(k)}$ and $Z_{0,i}^{(k)}$ for $i = 1, 2, \cdots, n$ and $k = 1, 2$ as $X_{0,i}^{(k)} = \sigma_{G,0}^{(k)} \sqrt{(Y_{0,i}^{(k)})^2 + (Z_{0,i}^{(k)})^2}$, where $Y_{0,i}^{(k)}$ and $Z_{0,i}^{(k)}$ are generated by the recursions

$$\begin{aligned} Y_{0,1}^{(k)} &= V_{0,1}^{(k)} \\ Y_{0,i}^{(k)} &= \rho_{G,0}^{(k)} Y_{0,i-1}^{(k)} + \sqrt{1 - [\rho_{G,0}^{(k)}]^2} V_{0,i}^{(k)};\ i > 1 \end{aligned}$$

and

$$\begin{aligned} Z_{0,1}^{(k)} &= U_{0,1}^{(k)} \\ Z_{0,i}^{(k)} &= \rho_{G,0}^{(k)} Z_{0,i-1}^{(k)} + \sqrt{1 - [\rho_{G,0}^{(k)}]^2} U_{0,i}^{(k)};\ i > 1 \end{aligned}$$

where $-1 \leq \rho_{G,0}^{(k)} \leq 1$ are correlation coefficients and, for sensor k, $V_{0,i}^{(k)}$, $U_{0,i}^{(k)}$ $(i = 1, 2, \cdots)$ are two i.i.d. standard Gaussian $(\mathcal{N}(0,1))$ sequences. In the case of dependence across time only, $(V_{0,i}^{(1)}, V_{0,i}^{(2)})$ are generated independently and so are $(U_{0,i}^{(1)}, U_{0,i}^{(2)})$). In the case of dependence across time and sensors, they are generated in a coupled manner as

$$V_{0,i}^{(2)} = \rho_{G,0}^{(c)} V_{0,i}^{(1)} + \sqrt{1 - [\rho_{G,0}^{(c)}]^2} W_{0,i} \quad \text{and} \quad U_{0,i}^{(2)} = \rho_{G,0}^{(c)} U_{0,i}^{(1)} + \sqrt{1 - [\rho_{G,0}^{(c)}]^2} \tilde{W}_{0,i}$$

where $W_{0,i}$ and $\tilde{W}_{0,i}$ are two i.i.d. standard Gaussian sequences independent of each other, and $-1 \leq \rho_{G,0}^{(c)} \leq 1$ is a correlation parameter; $U_{0,i}^{(1)}$, $V_{0,i}^{(1)}$, $W_{0,i}$, and $\tilde{W}_{0,i}$ are mutually independent.

To simplify the presentation of the simulation results, we consider a two-sensor/fusion-center system with sensor observations that have identical statistics. i.e., the first- and second-order pdfs, respectively, of the sensor observations are the same for the two sensors. This implies that the nonlinearities and quantizers employed by the sensors are the same.

1. The Case of Dependence Across Time

Example 1: This example pertains to the Scheme 1 described in Subsection III.1: fusion of unquantized sensor test statistics. The values of parameters are selected as $\sigma_{G,0}^{(1)} = \sigma_{G,0}^{(2)} = 2.0$ and $\rho_{G,0}^{(1)} = \rho_{G,0}^{(2)} = 0.9261$, under H_0; $\sigma_{G,1}^{(1)} = \sigma_{G,1}^{(2)} = 0.4915$, $\mu_{G,1}^{(1)} = \mu_{G,1}^{(2)} = 0.7982$ and $\rho_{G,1}^{(1)} = \rho_{G,1}^{(2)} = 0.9924$, under H_1. In the kernel of the integral equations the infinite sums (for the ρ-mixing case treated here) were truncated to $m_{1,1} = m_{2,1} = 300$ terms under hypothesis H_1 and $m_{1,0} = m_{2,0} = 30$ terms under hypothesis H_0, for the two sensors. The above values of the parameters are characteristic of cases of practical interest in radar discrimination. According to these values of the parameters, the first- and second-order moments of the observations under H_0 are the same as the corresponding ones under H_1, which makes the discrimination problem particularly difficult. Since, under each hypothesis, the univariate and bivariate densities of the observations of the two sensors are identical, the two optimal nonlinearities have the same form $g_{opt}(x)$. In Fig. 1.a, g_{opt} and the nonlinearity of $g_{iid}(x) = \ln[f_1(x)/f_0(x)]$ (the i.i.d likelihood ratio function), $(f_i(x),;\ i = 0, 1$, being the univariate densities under $H_i)$ are presented. In Fig. 1.b, we present the receiver operating characteristics (ROCs) of this example, which show that the performance of $\bar{g}_{opt}$ is superior to that of $\bar{g}_{iid}$.

Example 2: This example pertains to Scheme 2 of Subsection III.2: fusion of suboptimally quantized sensor observations. The revalent parameter values are the same as the ones in Example 1. The corresponding suboptimal breakpoints and quantization levels for both quantizers q_{iid}

and q_{iid}, which are induced from g_{opt} and g_{iid} of Example 1, respectively, are given in Table 1; the resulted ROCs are drawn in Fig. 2. From the ROCs in Fig. 2, we see that the performance of the quantizer q_{opt} is similar to the one of q_{iid} in this specific example; this is due to the low number of quantization levels $M_1 = M_2 = 8$ employed, as this number increases, considerable improvement in the performance can be achieved.

Example 3: This example pertains to Scheme 3 of Subsection III.3: fusion of optimally quantized sensor observations. We use the same parameter values as in Example 1. The optimal quantizer Q_{opt} (breakpoints and levels) is shown in Table 2, where the quantizer Q_{iid} obtained by ignoring the dependence is also included. The resulting ROCs are drawn in Fig. 3, which indicates that the scheme employing the optimal quantizer provides superior performance to the one which ignores the dependence.

Example 4: This example pertains to Scheme 4 of Subsection III.4: fusion of binary sensor decisions. The parameter values are set to be the same as in Example 1. In Fig. 4, we compare the ROCs using the nonlinearities g_{opt} and g_{iid} given in Fig. 1a. From Fig. 4, we conclude that the scheme employing g_{opt} yields a performance superior to the one which ignores the dependence.

Then we compare the ROCs of the optimal nonlinearities or quantizers for the above four schemes in Fig. 4a, where we see that the performance of Scheme 1 is best (as intuitively expected), Scheme 4 follows, then Scheme 3, and the performance of Scheme 2 is the worst. The two schemes employing quantized observations did not perform that well for this small number of quantization levels (3-bit or 8-level quantizers).

2. The Case of Dependence Across Time and Sensors

Example 5: This example pertains to the same scheme as in Example 1. The values of the parameters $\sigma_{G,i}^{(k)}$, $\mu_{G,i}^{(k)}$ and $\rho_{G,i}^{(k)}$ are assigned the same as in Example 1, but here $m_{1,1} = m_{2,1} = 600$ and $m_{1,0} = m_{2,0} = 60$ are used. In addition, we choose $\rho_{G,0}^{(c)} = 0.90$ and $\rho_{G,1}^{(c)} = 0.99$. In Fig. 5a, we present the plots of the nonlinearities $\bar{g}_{opt}$ (which takes into account the dependence across time and sensors), g_{opt} (which ignores the dependence across sensors), and g_{iid} (which ignores the dependence across time and sensors); g_{opt} and g_{iid} have the same forms as in Example 1 since the relevant parameter values are the same. In Fig. 5b, we draw the ROCs of the unquantized fusion schemes employing these three nonlinearities. From this figure, we see that the performances of $\bar{g}_{opt}$ is superior to that of g_{opt}, which is in turn superior to that of g_{iid}.

Example 6: This example pertains to the same scheme as in Example 2. The parameter values are the same as in Example 5. The suboptimal quantizer $\bar{q}_{iid}$ (breakpoints and levels), which is induced from $\bar{g}_{opt}$ of Example 5, is given in Table 3. In Fig. 6, we draw the ROCs which use the three quantizers $\bar{q}_{opt}$, q_{opt} and q_{iid}; where q_{opt} and q_{iid} are obtained from the g_{opt} and g_{iid} as in Example 2. From Fig. 6, we notice that the performance of the suboptimal quantizers $\bar{q}_{opt}$ is superior to that of q_{opt}, which in turn is superior to that of q_{iid}.

Example 7: This example pertains to the same scheme as in Example 3. We use the same parameter values as in Example 5. In Table 4, the quantizer $\bar{Q}$ (breakpoints and levels) resulted from numerical optimization is given. In Fig. 7, we then present the ROCs for the schemes employing $\bar{Q}_{opt}$, Q_{opt} and Q_{iid}, respectively; Q_{opt} and Q_{iid} are as defined in Example 3. From Fig. 7, we observe that the performance of $\bar{Q}_{opt}$ is superior to that of Q_{opt} and Q_{iid} and the performance of Q_{opt} is superior to that of Q_{iid}.

Example 8: This example pertains to the same scheme as in Example 4. The parameter values are set to be the same as in Example 5. In Fig. 8, we present the ROCs which uses g_{opt} and g_{iid} (of Example 1) in this case. From this figure, we see that the performance of g_{opt} is better than that of g_{iid} for dependent observations across time and sensors.

In Fig. 8a, we provide the same comparison as the one presented in Fig. 4a. From this figure, we conclude that in this case, the performance of Scheme 1 is best, Scheme 4 follows, then Scheme 3, and, finally, Scheme 2 which has the worst performance.

VI. EXTENSIONS AND CONCLUSIONS

We conclude this paper by discussing some extensions of the work presented in the previous sections. First we consider minimax robust discrimination for an environment with uncertainty in the sensor observations, and then a multi-sensor system with more than two sensors. Finally, conclusions are drawn from this work.

1. Robustness

In the derivation of the design criteria and the determination of the optimal nonlinearities and/or quantizers for the four schemes in Sections III and IV, the univariate and bivariate densities of the sensor observations under the two hypotheses were assumed known a priori. However, this assumption is not valid in many practical situations. Thus the robust design of the nonlinearities and quantizers based on models of uncertainty in the univariate and bivariate pdfs is of interest. A survey of robust single-sensor detection techniques is given in [20]. Here we

pursue maximin robust designs according to which we solve optimization problems (involving the design criteria of Sections III and IV) of the form

$$\max_{g_k} \min_{f_{k,i}, f_{k,i}^{(j)}} \bar{J}_k(g_k, f_{k,i}, f_{k,i}^{(j)}) \tag{126}$$

where $i = 0, 1$, $j = 1, 2, \cdots, m_{k,i}$ for $f_{k,i}^{(j)}$, and $k = 1, 2$, for the case of dependence across time only; and

$$\max_{g} \min_{f_i, f_i^{(j)}, \tilde{f}_i^{(j)}} \bar{J}_{\text{mm}}(g, f_i, f_i^{(j)}, \tilde{f}_i^{(j)}) \tag{127}$$

where $i = 0, 1$, $j = 1, 2, \cdots, m$ for $f_i^{(j)}$, and $j = 0, 1, \cdots, \tilde{m}$ for $\tilde{f}_i^{(j)}$, for the case of dependence across time and sensors. In these formulations the univariate and bivariate pdfs belong to specific uncertainty classes as discussed in the following paragraph. The above maximin problems were formulated for the nonlinearities only, when quantizers are involved instead of or in addition to the nonlinearities, we maximize with respect to q_k or (g_k, q_k), respectively.

In the previous sections we derived two basic distinct forms of objective functions for the design of optimal nonlinearities and quantizers: $(\mu_{k,1} - \mu_{k,0})^2/(\sigma_{k,0}^2 + \sigma_{k,1}^2)$ $(k = 1, 2)$ and $(\mu_1 - \mu_0)^2/[(1+\rho_0)\sigma_0^2 + (1+\rho_1)\sigma_1^2]$. The minimax robustness for the objective function $(\mu_{k,1} - \mu_{k,0})^2/(\sigma_{k,0}^2 + \sigma_{k,1}^2)$ $(k = 1, 2)$ has been addressed in our work of [13] in a single-sensor context. In [13] we modeled the uncertainty in the univariate pdfs by Huber-Strassen capacity classes which include the popular ϵ-contaminated, total variation, band, and p-point classes (see [13] and [18]). For the bivariate pdf under hypothesis H_i $(i = 0, 1)$ the following uncertainty class was considered in [13]

$$\sup_{g_k} \frac{|cov_i\{g_k(X_1^{(k)}), g_k(X_{j+1}^{(k)})\}|}{\sqrt{var_i\{g_k(X_1^{(k)})\} var_i\{g_k(X_{j+1}^{(k)})\}}} \leq r_{k,i}^{(j)} \quad k = 1, 2;\ j = 1, 2, \cdots, m_{k,i} \tag{128}$$

for all nonlinearities g_k. This class was first introduced in [19] for robust single-sensor weak-signal detection and estimation. Thus for the design criterion $(\mu_{k,1} - \mu_{k,0})^2/(\sigma_{k,0}^2 + \sigma_{k,1}^2)(k = 1, 2)$ used for Schemes 1 and 2 in the case of dependence across time (Subsections III.1 and III.2), the least-favorable univariate and bivariate densities derived in [13] can be applied directly. For Scheme 4, since the optimal nonlinearities in both Subsections III.4 and IV.4 are determined by maximizing the same objective function (the one described above), the maximin robust results of [13] are still applicable. Unfortunately, for Scheme 3 in both cases of dependence across time as well as across time and sensors (Subsections III.3 and IV.3), we were unable to derive explicit

maximin robust results due to the lack of closed form expressions for the optimal quantization break-points; further investigation is necessary here.

For the other objective function $(\mu_1 - \mu_0)^2/[(1+\rho_0)\sigma_0^2 + (1+\rho_1)\sigma_1^2]$ which is employed as the design criterion for Schemes 1 and 2 in the case of dependence across time and sensors (Subsections IV.1 and IV.2), we can exploit the similarity with the the objective function $(\mu_1 - \mu_0)^2/(\sigma_0^2 + \sigma_1^2)$ of [13] ,and modify the formulation and results of [13] to obtain the desirable maximin robust design. In particular, we use the same uncertainty classes for the univariate pdfs as in [13], i.e., the Huber-Strassen classes mentioned above, and for the bivariate pdfs we require that besides belonging to the class satisfying (128) (for symmetric conditions) they also satisfy

$$\sup_g \frac{|cov_i\{g(X_1^{(1)}), g(X_{j+1}^{(2)})\}|}{\sqrt{var_i\{g(X_1^{(1)})\} var_i\{g(X_{j+1}^{(2)})\}}} \leq \tilde{r}_i^{(j)} \quad j = 0, 1, \cdots, \tilde{m}_i \tag{129}$$

under hypothesis H_i $(i = 0, 1)$ for all nonlinearities g. Consequently, under the stationarity assumption we can derive that

$$\begin{aligned} \sup_g\{(1+\rho_0)\sigma_0^2 + (1+\rho_1)\sigma_1^2\} \quad \leq \quad & \left[1 + 2\sum_{j=1}^{m_0} r_0^{(j)} + \tilde{r}_0^{(0)} + 2\sum_{j=1}^{\tilde{m}_0} \tilde{r}_0^{(j)}\right] \cdot var_0\{g\} \\ & + \left[1 + 2\sum_{j=1}^{m_1} r_1^{(j)} + \tilde{r}_1^{(0)} + 2\sum_{j=1}^{\tilde{m}_1} \tilde{r}_1^{(j)}\right] \cdot var_1\{g\} \end{aligned} \tag{130}$$

which is the generalization of Eq. (8) in theorem 2 of [19]. Following procedures similar to lemmas 1-3 of [19] we can construct least-favorable bivariate densities similar to those given in [19]. Then we follow the procedure of [13] to derive the desirable maximin robust nonlinearities.

2. Systems with More than Two Sensors

The results for Schemes 1-4 of Section III in the case of dependence across time can be directly extended to configurations with more than two sensors. In a K-sensor environment $(K > 2)$, it is straightforward to show that the same design criteria $\bar{J}_k$ $(k = 1, 2, \cdots, K)$ are involved, because the lower bound on the asymptotic rates of the error probabilities of the fusion center is additive in $\bar{J}_k$. Thus the optimal nonlinearities g_k obtained in Section III for Schemes 1-4 for $k = 1, 2$ are still applicable to the case $k = 1, 2, \cdots, K$ and are characterized by the optimization problem (57) with $\bar{J}_k$ given by (55) for sensor k.

For the case of dependence across time and sensors, however, even under the symmetric conditions (identical sensor observation statistics for different sensors) imposed in Section IV,

coupled nonlinear integral equations are involved in the optimization whose solution requires numerical search. No explicit analytic solutions can be derived in this case.

3. Conclusions

In this paper we have designed, analyzed, and simulated the performance of new multisensor fusion and quantization schemes for discrimination (nonlocal detection) between different hypotheses on the basis of stationary dependent observations. These schemes employ memoryless nonlinearities and take partial advantage of the dependence in the observations across time and/or sensors. For large sample sizes, our schemes are optimal within the class of sensors using memoryless nonlinearities and require knowledge of the univariate and bivariate pdfs of sensor observations. Four different schemes with varying degrees of optimality in the quantization and fusion procedures were introduced, analyzed, and simulated.

The dependence in the sensor observations was characterized by one of the mixing types: m-dependent, ϕ-mixing, or ρ-mixing. The performance of the two-sensor/fusion-center configuration was measured by the error probabilities of the fusion center under the two hypotheses.

For three of the quantization/fusion schemes treated in this paper the generalized signal-to-noise ratio (GSNR) turned out to be the appropriate criterion as a byproduct of the Neyman-Pearson formulation on the fusion center error probabilities. For the fourth scheme the asymptotic rate of the error probabilities of the fusion center was derived as the performance measure and this also resulted in the GSNR as the criterion for the final optimization. For the case of dependence across time and sensors a modified GSNR which involves the correlation coefficients of the test statistics of the two sensors under the two hypotheses was derived.

Optimizing with respect to the nonlinearities and/or quantizers in the different cases has led to uncoupled or coupled linear integral equations involving the univariate and bivariate pdfs of the sensor observations. In the case of dependence across time only the resulting uncoupled linear integral equations are readily solved. In the case of dependence across time and sensors the resulting coupled nonlinear integral equations are difficult to solve; however, for the special case of identical sensor statistics they reduce to a single linear integral equation which is readily solved.

Comparison of the performance of the various schemes via simulation establishes that, as expected, fusing the unquantized test statistics is superior to all schemes, and fusing binary (hard) decisions is superior to fusing quantized observations (at least for small numbers of

quantization levels). Optimal quantization is superior to uniformly-spaced discretization of the continuous nonlinearities. Moreover, two-sensor discrimination schemes which take into consideration the dependence across time and sensors outperform the corresponding schemes which ignore the dependence across sensors, and these schemes outperform by a large margin the corresponding schemes which ignore the dependence across time and sensors.

Maximin robustification of most of the proposed two-sensor discrimination schemes against uncertainties in the univariate and bivariate pdfs of the observation processes can be carried out by applying existing results from single-sensor systems to the case of dependence across time and the one of symmetric densities for dependence across time and sensors. In addition, in the case of dependence across time, the proposed discrimination schemes can be directly extended to configurations involving more than two sensors. In the case of dependence across time and sensors, the basic methodology is still applicable, but the solution of nonlinear integral equations is required for the optimal nonlinearities and/or quantizers.

Appendix A: Proofs of Lemma 1 and Theorem 2

Proof of Lemma 1: By using arguments similar to the ones used in the proof of theorem 2 in [12], we write the following Taylor series expansion for $\bar{b}_i(s)$

$$\bar{b}_i(s) = \lim_{n\to\infty} \frac{1}{n} E_i[\ln L_n]s + o(s)$$

where $o(s)$ represents high-order terms with the property $o(s)/s \to 0$ as $s \to 0$. Thus, as $n \to \infty$

$$\bar{b}_i(s) \to \varphi_i s$$

according to (12) and (13). Therefore, from the definitions of B_s, $I_i(z)$ and $\tilde{s}_i$

$$I_i(z) = \sup_{s\in B}\{sz - \bar{b}_i(s)\} = \sup_{s\in B}\{(z-\varphi_i)s\} = (z-\varphi_i)\tilde{s}_i.$$

Moreover, since $I_i(z)$ is the supreme of $sz - \bar{b}_i(s)$ on the set B_s, we can always choose the sign of $\tilde{s}_i$ such that $I_i(z)$ is positive for any z.

Proof of Theorem 2: We first show the proof of the asymptotic rate for $P_0(\ln L_n > n\eta)$. Let $z = \frac{1}{n}\ln L_n$ and define the sets G and $\tilde{G}$ as

$$G = \{\ln L_n \in R : \ln L_n > n\eta\} = \{z \in R : z > \eta\}$$

and

$$\tilde{G} = \{\ln L_n \in R : \ln L_n \geq n\eta\} = \{z \in R : z \geq \eta\},$$

in addition, let $\bar{G}$ be the closure of G. Then from the definition of $\bar{I}_i(\bar{G})$, we have $\bar{I}_i(G) = \bar{I}_i(\tilde{G}) = \bar{I}_i(\bar{G})$. We also have $P_0(G) \leq P_0(\tilde{G}) \leq P_0(\bar{G})$ since $G \subset \tilde{G} \subset \bar{G}$. Then from Theorem 1, we have

$$\begin{aligned}
& \liminf_{n\to\infty} -\frac{1}{n}\ln P_0(\ln L_n > n\eta) = \liminf_{n\to\infty} -\frac{1}{n}\ln P_0(G) \\
\geq\ & \liminf_{n\to\infty} -\frac{1}{n}\ln P_0(\bar{G}) \\
\geq\ & \bar{I}_i(\bar{G}) = \bar{I}_i(G) \\
\geq\ & \limsup_{n\to\infty} -\frac{1}{n}\ln P_0(G) \\
=\ & \limsup_{n\to\infty} -\frac{1}{n}\ln P_0(\ln L_n > n\eta)
\end{aligned}$$

since $\{\ln L_n : \ln L_n > n\eta\}$ is an open subset of $\mathbf{R}$. On the other hand,

$$\limsup_{n\to\infty} -\frac{1}{n}\ln P_0(\ln L_n > n\eta) \geq \liminf_{n\to\infty} -\frac{1}{n}\ln P_0(\ln L_n > n\eta).$$

Consequently,

$$\lim_{n\to\infty} -\frac{1}{n}\ln P_0(\ln L_n > n\eta) = \bar{I}_i(z > \eta).$$

The above results were stated and used in [17]; a similar formulation appeared in [12]. We show them here for completeness. Therefore,

$$\lim_{n\to\infty} -\frac{1}{n}\ln P_0(\ln L_n > n\eta) = \bar{I}_i(z > \eta) \geq I_i(z > \eta) = \inf_{\{z>\eta\}} I_i(z) = \inf_{\{z>\eta\}} (z - \varphi_0)\tilde{s}_0$$

where the last inequality is from (26) and the last equality is from Lemma 1 for $i = 0$. Under the consistency condition $\varphi_0 \leq \eta \leq \varphi_1$, we have

$$\inf_{\{z>\eta\}} (z - \varphi_0)\tilde{s}_0 = (\eta - \varphi_0)\tilde{s}_0$$

where because $\eta - \varphi_0 \geq 0$, $\tilde{s}_0$ is set to be positive according to Lemma 1. Following steps similar to the above ones we can obtain

$$\lim_{n\to\infty} -\frac{1}{n}\ln P_1(\ln L_n \leq n\eta) \geq \inf_{\{z\leq\eta\}} (z - \varphi_1)\tilde{s}_1 = (\eta - \varphi_1)\tilde{s}_1$$

where because $\eta - \varphi_1 \leq 0$ under the consistent condition, $\tilde{s}_1$ is set to be negative. Equivalently, we have

$$\lim_{n\to\infty} -\frac{1}{n}\ln P_1(\ln L_n \leq n\eta) \geq (\varphi_1 - \eta)|\tilde{s}_1|.$$

Since both $\tilde{s}_i$ are small numbers near zero, we define

$$\tilde{s} = \min\{\tilde{s}_0, |\tilde{s}_1|\}$$

and thus

$$\lim_{n\to\infty} -\frac{1}{n}\ln P_0(\ln L_n > n\eta) \geq (\eta - \varphi_0)\tilde{s}$$

and

$$\lim_{n\to\infty} -\frac{1}{n}\ln P_1(\ln L_n \leq n\eta) \geq (\varphi_1 - \eta)\tilde{s}$$

which are the two inequalities in (27). Finally, (29) with I_{min} defined by (28) results from (15)

Appendix B: Maximization of (53)

It is easily to check that J_k is invariant under the scaling of g_k. Consequently, the maximization problem characterized by (53) is equivalent to maximizing

$$\bar{H}(g_k) = \mu_{k,1} - \mu_{k,0} + \lambda\left[\frac{1}{\sigma^2_{k,0}} + \frac{1}{\sigma^2_{k,1}}\right]$$

where λ is the Lagrange multipler. The necessary condition of the maximization of $\bar{H}(g_k)$ is

$$\frac{\partial}{\partial \epsilon}\bar{H}(g_k + \epsilon\delta g_k)|_{\epsilon=0} = 0$$

where

$$\begin{aligned}
\frac{\partial \bar{H}(g_k + \epsilon\delta g_k)}{\partial \epsilon}|_{\epsilon=0} =& \int \delta g_k(x)dx \{f_{k,1}(x) - f_{k,0}(x) \\
& -2\lambda \left\{ \frac{1}{\sigma^4_{k,0}} \left[g_k(x)f_{k,0}(x) + \int [2\sum_{j=1}^{m_{k,0}} [f^{(j)}_{k,0}(x,y) - f_{k,0}(x)f_{k,0}(y)] - f_{k,0}(x)f_{k,0}(y)]g_k(y)dy \right] \right. \\
& \left. + \frac{1}{\sigma^4_{k,1}} \left[g_k(x)f_{k,1}(x) + \int [2\sum_{j=1}^{m_{k,1}} [f^{(j)}_{k,1}(x,y) - f_{k,1}(x)f_{k,1}(y)] - f_{k,1}(x)f_{k,1}(y)]g_k(y)dy \right] \right\} \\
=& \int \left\{ f_{k,1}(x) - f_{k,0}(x) - 2\lambda \left\{ \left[\frac{f_{k,1}(x)}{\sigma^4_{k,1}} + \frac{f_{k,0}(x)}{\sigma^4_{k,0}} \right] g_k(x) \right. \right. \\
& \left. \left. + \int \left[\frac{K_{k,1}(x,y)}{\sigma^4_{k,1}} + \frac{K_{k,0}(x,y)}{\sigma^4_{k,0}} \right] g_k(y)dy \right\} \right\} \delta g_k(x)dx
\end{aligned}$$

with $K_{k,i}(x,y)$ defined by

$$K_{k,i}(x,y) = 2\sum_{j=1}^{m_{k,i}} [f^{(j)}_{k,i}(x,y) - f_{k,i}(x)f_{k,i}(y)] - f_{k,i}(x)f_{k,i}(y).$$

This leads to the following nonlinear integral equation

$$\frac{f_{k,1}(x) - f_{k,0}(x)}{\frac{f_{k,1}(x)}{\sigma^4_{k,1}(g_k)} + \frac{f_{k,0}(x)}{\sigma^4_{k,0}(g_k)}} = g_k(x) + \int \left[\frac{K_{k,1}(x,y)}{\sigma^4_{k,1}(g_k)} + \frac{K_{k,0}(x,y)}{\sigma^4_{k,0}(g_k)} \right] g_k(y)dy$$

Unfortunately, we could not obtain a sufficient condition for this optimization problem. Thus, although the above nonlinear integral equation can be solved via numerical techniques, we have no way of guaranteeing the optimality of the solution.

Appendix C: Maximization of (99)

We can easily verify that $\bar{J}$ is invariant under the scaling of g. By using similar arguments as in Appendix B, the maximization of (99) is equivalent to the maximization of

$$\tilde{H}(g) = \mu_1 - \mu_0 + \lambda[(1+\rho_1)\sigma_1 + (1+\rho_0)\sigma_0]$$

where λ is the Lagrange multiplier and, for $i = 0, 1$,

$$\begin{aligned}
\rho_i\sigma_i^2 &= E_i \left\{ \frac{[\sum_{i=1}^n g(X_i^{(1)}) - \mu_i]}{\sqrt{n}} \cdot \frac{[\sum_{i=1}^n g(X_i^{(2)}) - \mu_i]}{\sqrt{n}} \right\} \\
&= E_i\{g(X_1^{(1)})g(X_1^{(2)})\} - \mu_i^2 + 2\sum_{j=1}^{\tilde{m}_i} [E_i\{g(X_1^{(1)})g(X_{j+1}^{(2)})\} - \mu_i^2]
\end{aligned}$$

The necessary condition for optimal g is given by

$$\frac{\partial \tilde{H}(g+\epsilon\delta g)}{\partial \epsilon}|_{\epsilon=0} = 0$$

which leads to the linear integral equation

$$0 = f_1(x) - f_0(x) + 2\lambda g(x)[f_1(x) + f_0(x)]$$
$$+2\lambda \int \left[\sum_{i=0}^{1} \left\{ 2 \sum_{j=1}^{m_i} [f_i^{(j)}(x,y) - f_i(x) f_i(y)] + 2 \sum_{j=1}^{\bar{m}_i} [\tilde{f}_i^{(j)}(x,y) - f_i(x) f_i(y)] + [\tilde{f}_i^{(0)}(x,y) - 2 f_i(x) f_i(y)] \right\} \right] g(y) dy.$$

The sufficient condition is given by

$$\frac{\partial^2 \tilde{H}(g+\epsilon\delta g)}{\partial \epsilon^2}|_{\epsilon=0} < 0.$$

and is satisfied if $\lambda < 0$. Actually, we set $\lambda = -1/2$ in order to be consistent with the single-sensor discrimination formulation of [12]-[13]. Finally, the optimal g is obtained by solving the linear integral equation given by (100) with the kernel defined by (101).

References

[1] R. R. Tenny and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 501-510, July 1981.

[2] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-22, pp. 98-101, January 1986.

[3] A. R. Reibman and L. W. Nolte, "Optimal detection and performance of distributed sensor systems," *IEEE Trans. Aerosp. Electron. Systems,* vol. AES-23, pp. 644-653, Jan. 1987.

[4] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems,* vol.2, 1988.

[5] E. Geraniotis and Y. A. Chau, "Robust data fusion for multi-sensor detection systems," to appear in *IEEE Trans. Inform. Theory,* 1990.

[6] G. S. Lauer and N. R. Sandell, Jr., "Distributed detection with waveform observations: correlated observation processes", *Proceedings of the 1982 American Control Conference* pp. 812-817.

[7] J. J. Chao and C. C. Lee, "A distributed detection scheme based on soft local decisions," *Proceedings of 24th Annual Allerton Conference on Communication, Control, and Computing*, Oct. 1986.

[8] Y. A. Chau and E. Geraniotis, "Distributed detection from multiple sensors with correlated observations: A large deviations approach," submitted for publication to *IEEE Trans. Automatic Control*, 1989.

[9] Y. A. Chau and E. Geraniotis, "Memoryless quantization and fusion in multi-sensor systems for the detection of weak signals in dependent noise," submitted for publication to *IEEE Trans. Inform. Theory*, 1989.

[10] R. C. Bradley, "Basic properties of strong mixing conditions ," *Dependence in Probability and Statistics*, Birkhauser, 1985.

[11] M. Peligrad, "Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables (a survey)," *Dependence in Probability and Statistics*, Birkhauser, 1985.

[12] J. S. Sadowsky and J. Bucklew, "A nonlocal approach for asymptotic memoryless detection theory," *IEEE Trans. Inform. Theory,* vol. IT-32, pp.115-120, Jan. 1986.

[13] D. Sauder and E. Geraniotis, "Optimal and robust memoryless discrimination from dependent observations", submitted for publication to *IEEE Trans. Inform. Theory*, 1989.

[14] S. A. Kassam, "Optimum quantization for signal detection," *IEEE Trans. Commun.,* vol. COM-25, pp. 479-484, May 1977.

[15] H. V. Poor and J. B. Thomas,"Memoryless quantizer-detectors for constant signals in m-dependent noise", *IEEE Trans. Inform. Theory* , vol. IT-26, pp. 423-432, July 1980.

[16] W. A. Gardner, " A unifying view of second-order measures of quality for signal classification," *IEEE Trans. Commun.*, vol. COM-28, pp. 807-816, June 1980.

[17] R. S. Ellis, "Large deviations for a general class of random vectors," *Ann. Probab.,* vol. 12, pp. 1-12, 1976.

[18] S. A. Kassam and H. V. Poor,"Robust techniques for signal processing: A survey," *Proc. IEEE,* vol. 73, pp. 433-481, 1985.

[19] J. S. Sadowsky, "A maximum variance model for robust detection and estimation with dependent data," *IEEE Trans. Inform. Theory,* vol. IT-32, pp. 220-226, Mar. 1986.

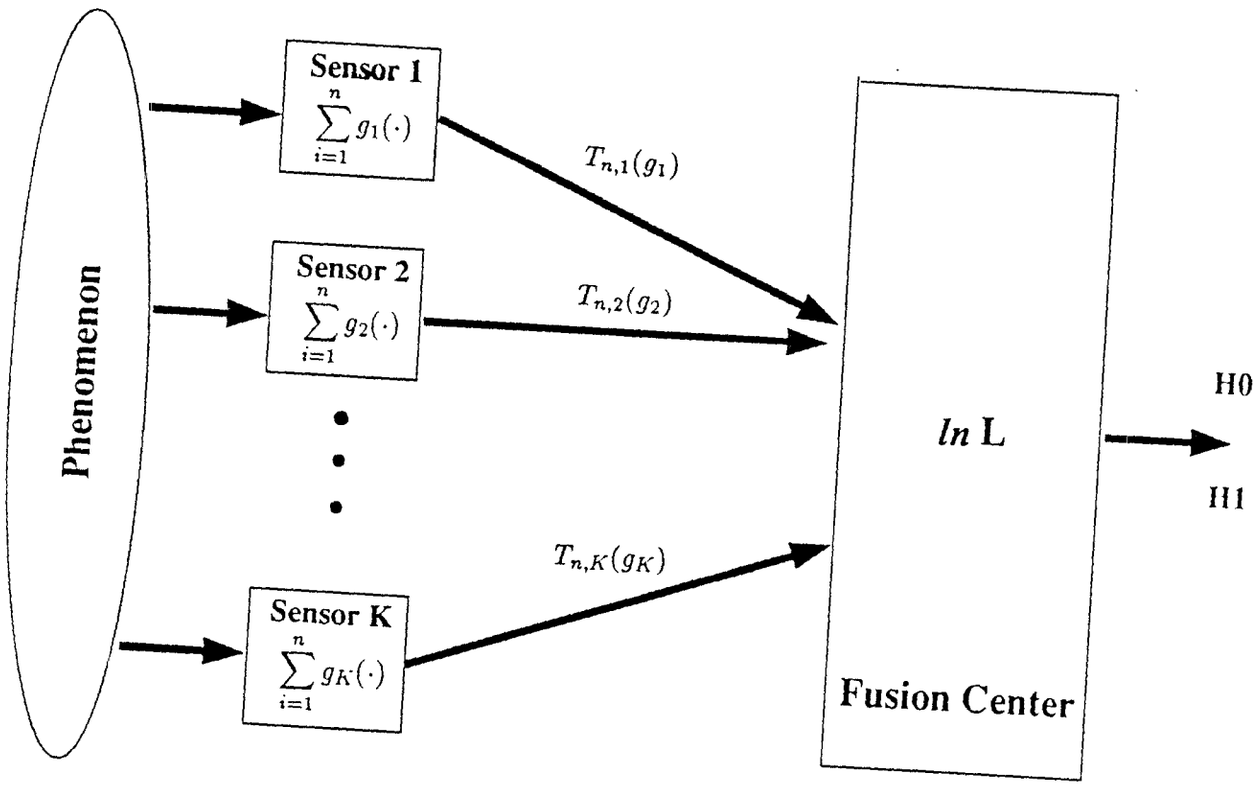


Fig. C.1 Fusion of Unquantized Sensor Test Statistics (Scheme 1)

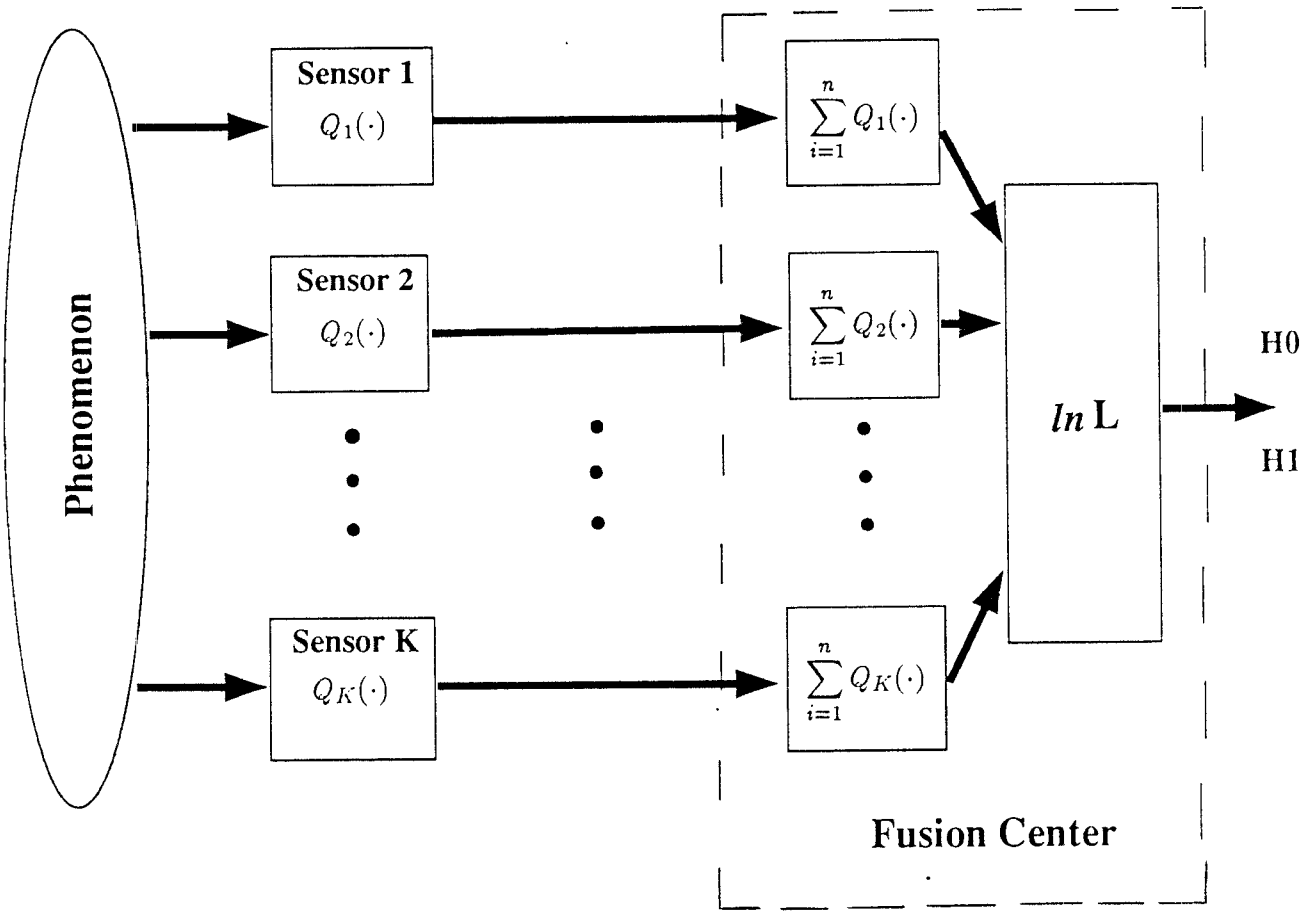


Fig. C.2 Fusion of Quantized Sensor Observations (Schemes2 and 3)

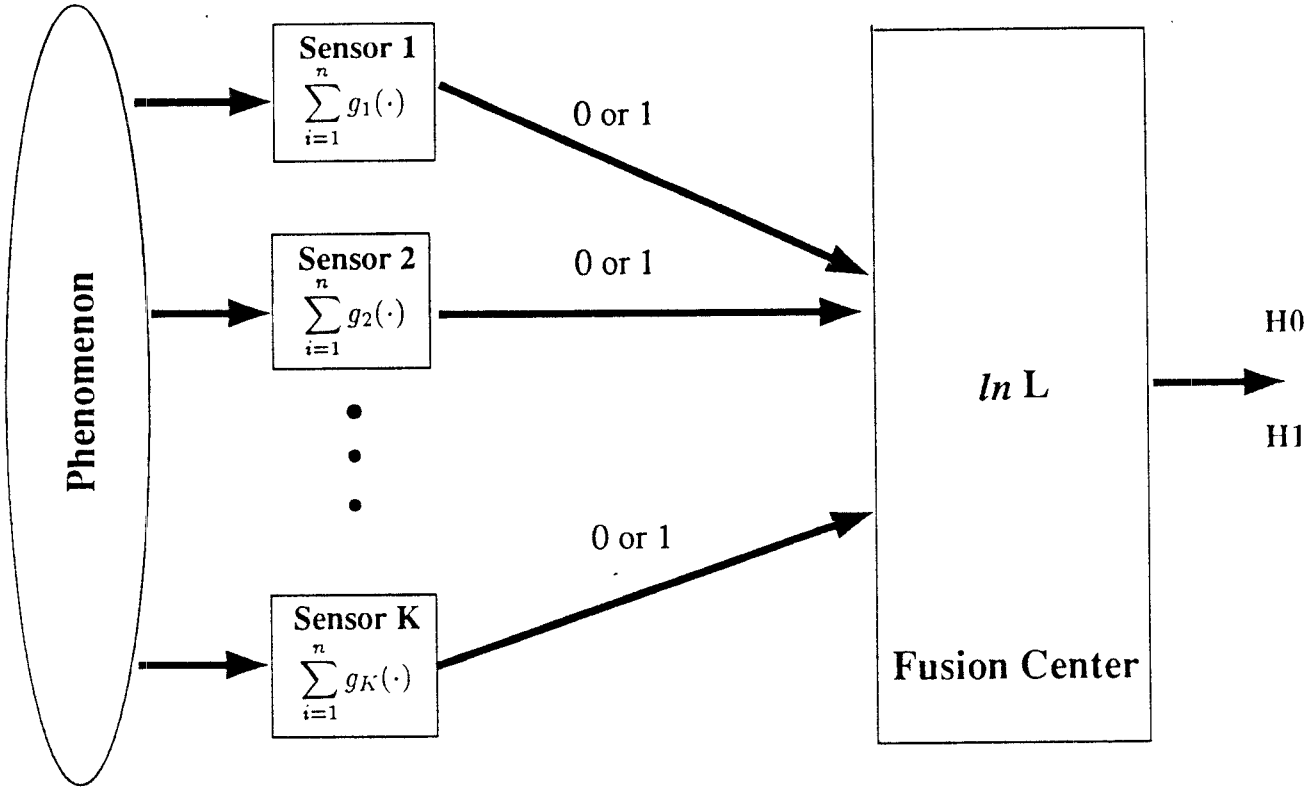


Fig. C.3 Fusion of Binary Sensor Decisions (Scheme 4)

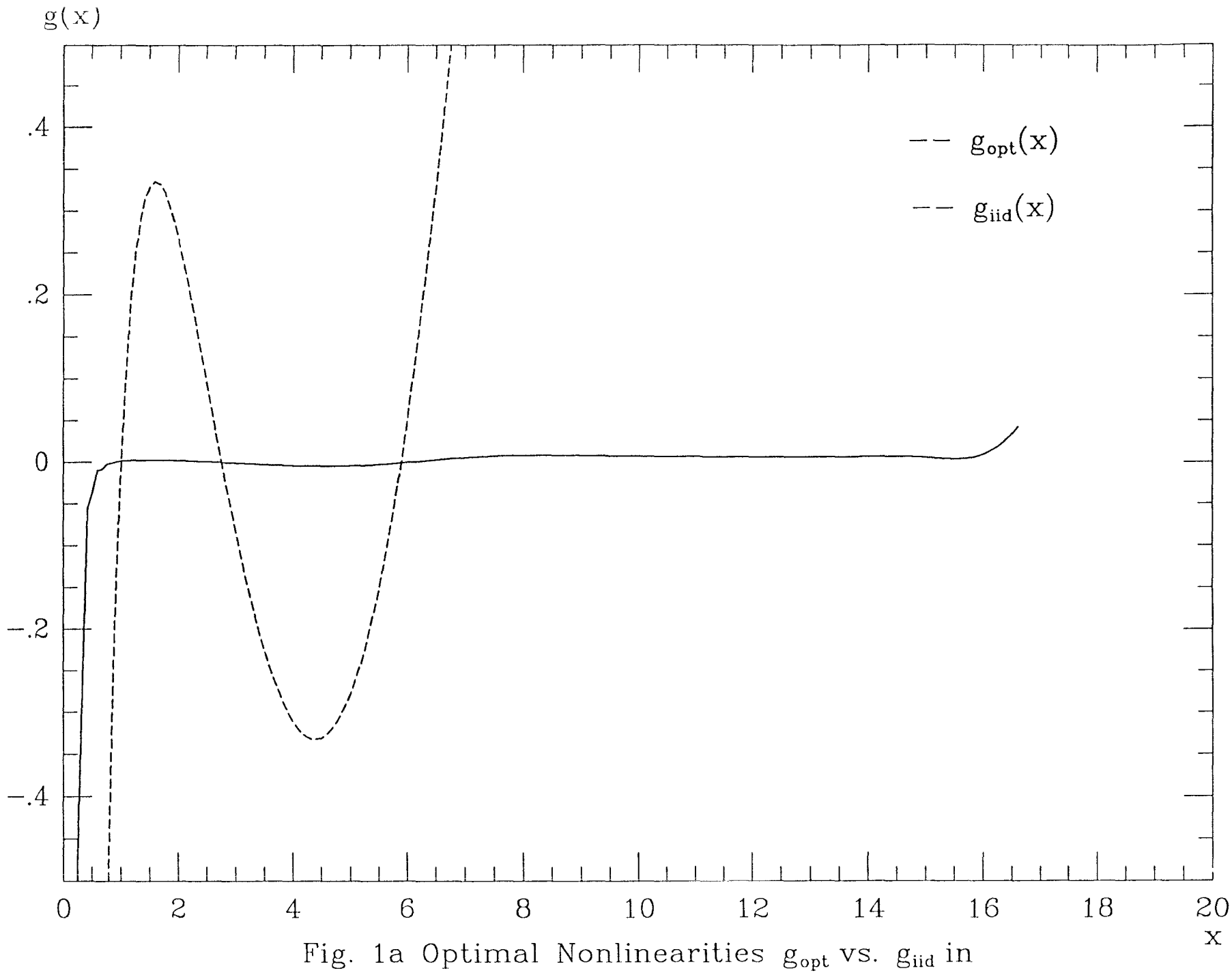


Fig. 1a Optimal Nonlinearities g_{opt} vs. g_{iid} in Examples 1, 4 and 8

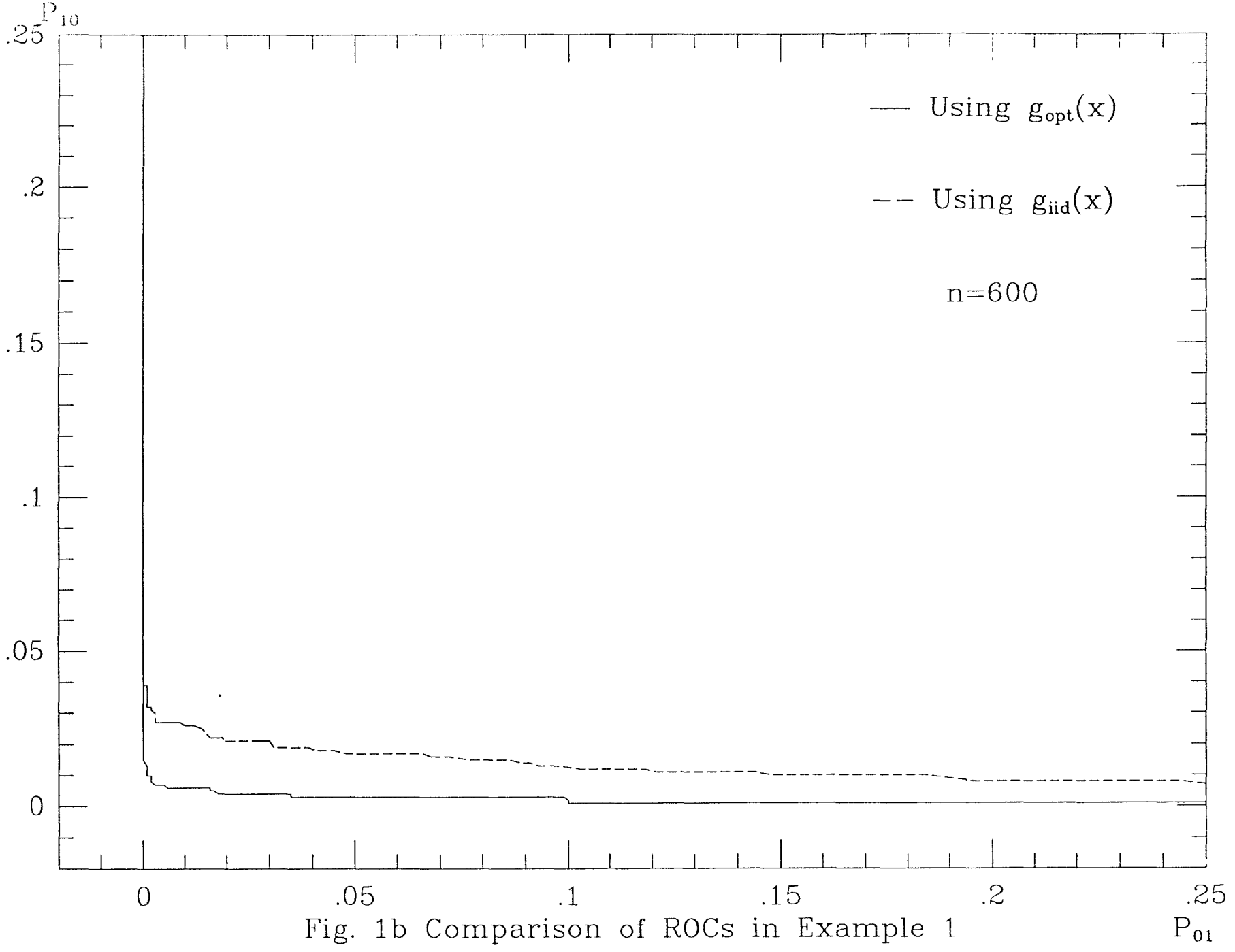


Fig. 1b Comparison of ROCs in Example 1

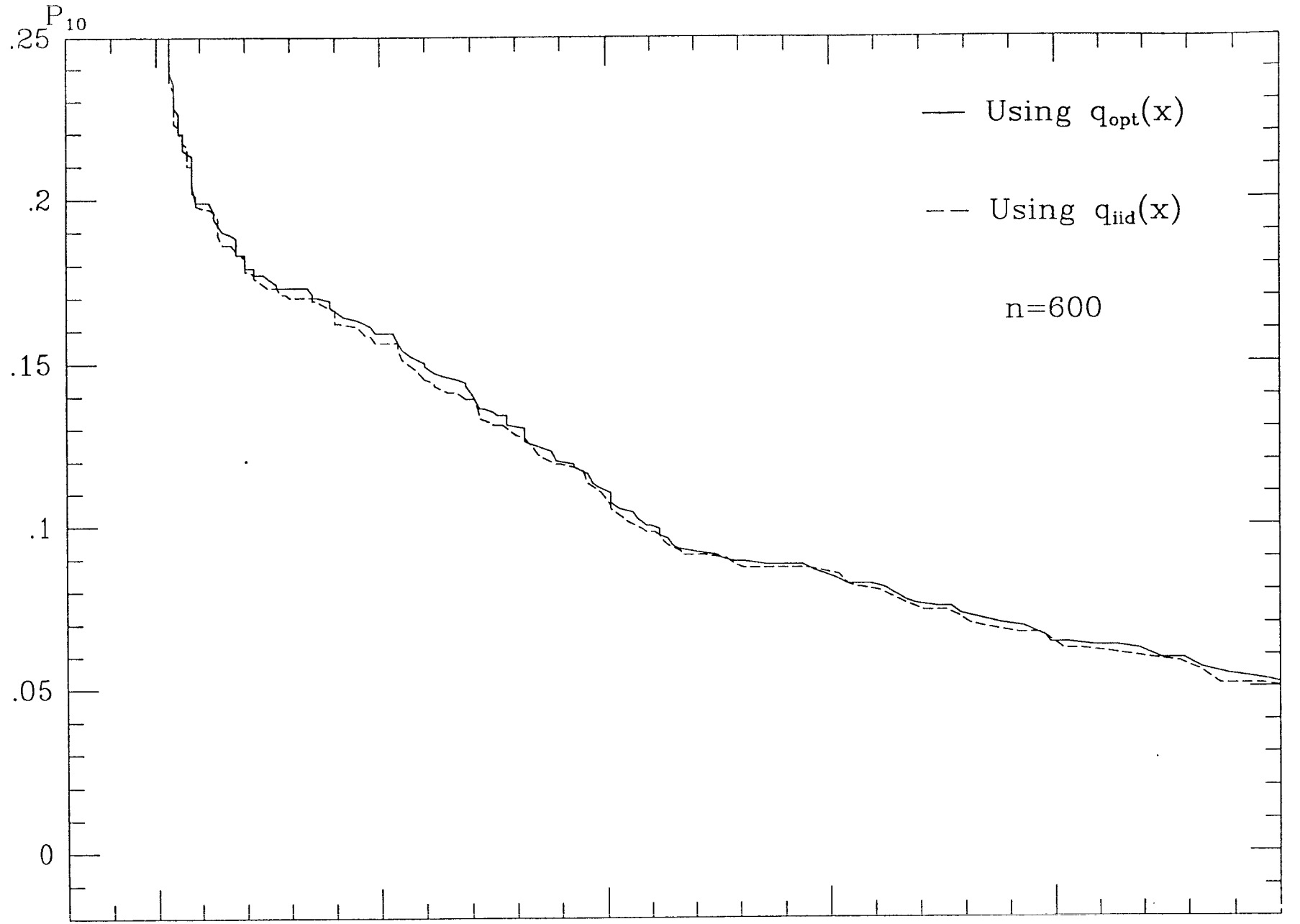


Fig. 2 Comparison of ROCs in Example 2

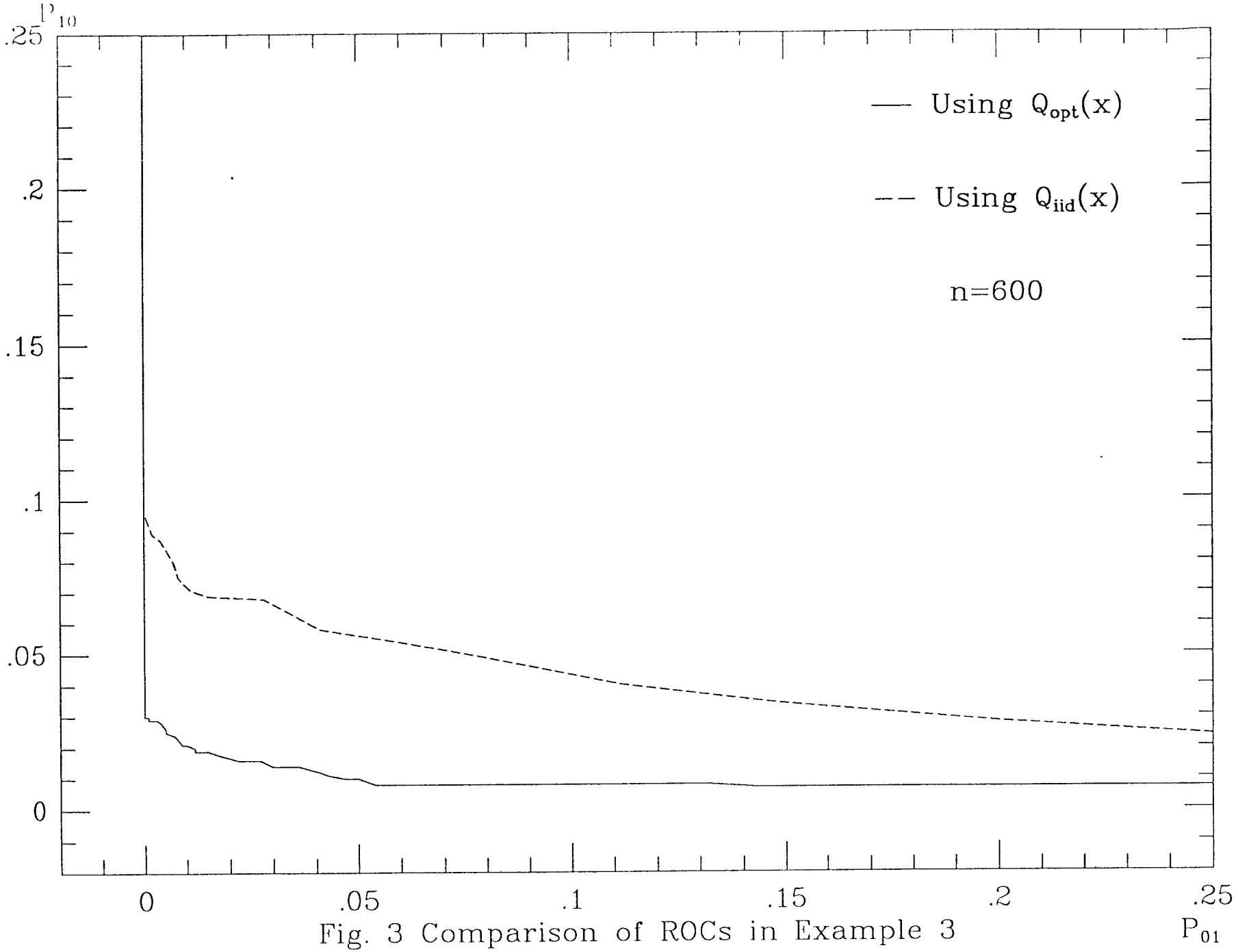


Fig. 3 Comparison of ROCs in Example 3

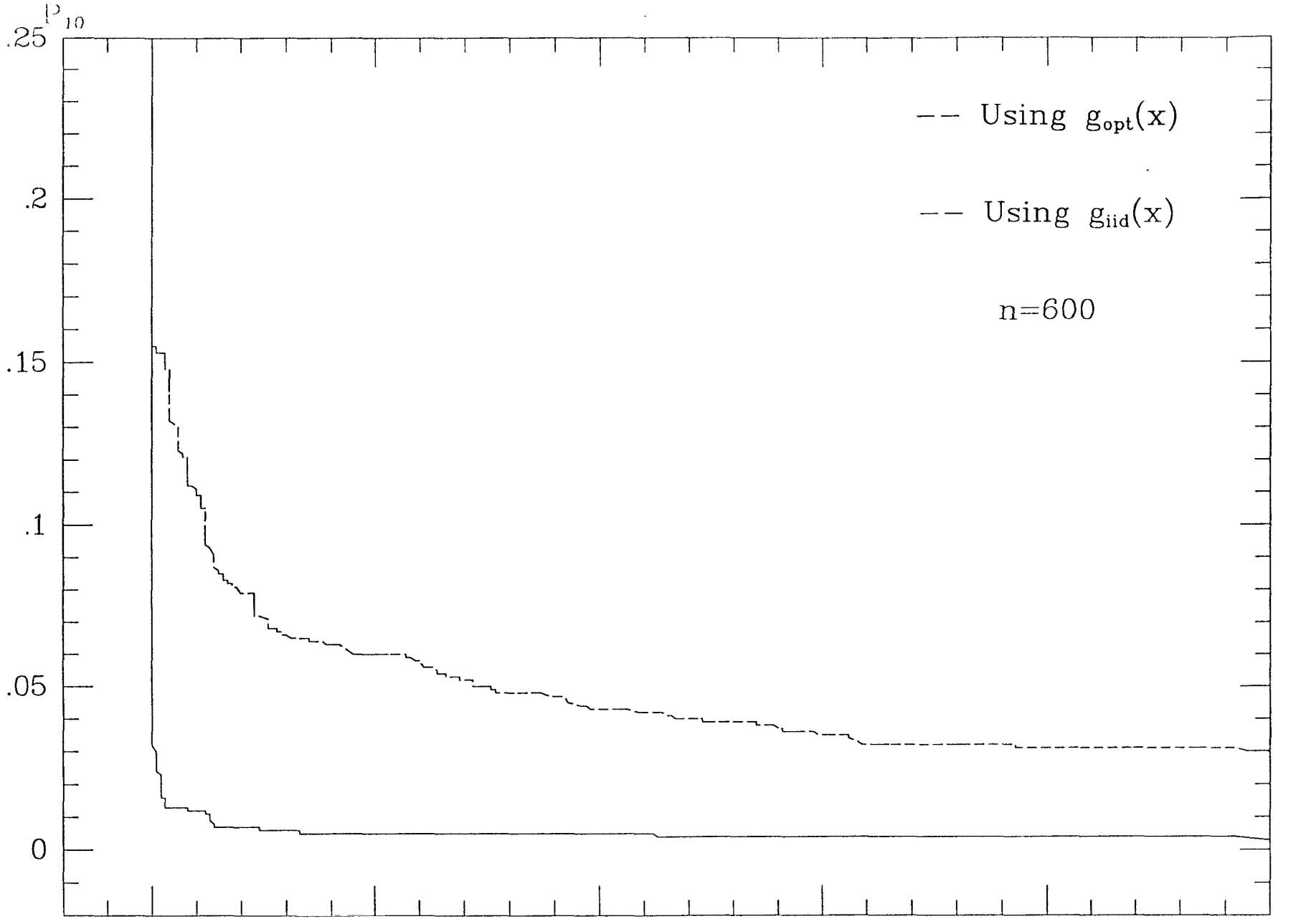


Fig. 4 Comparison of ROCs in Example 4

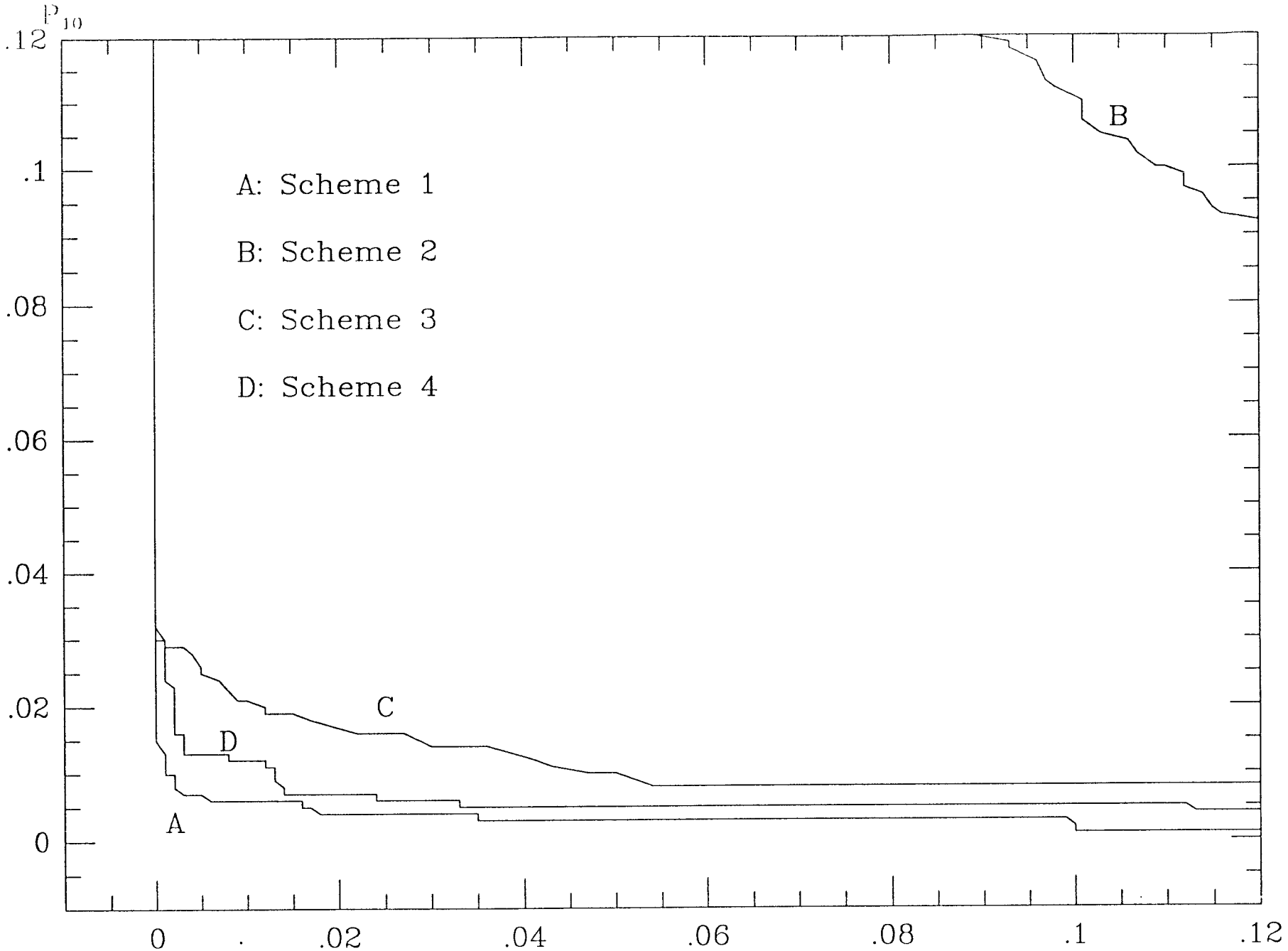


Fig. 4a Comparison of ROCs for Schemes 1–4 in the Case of Dependence Across Time

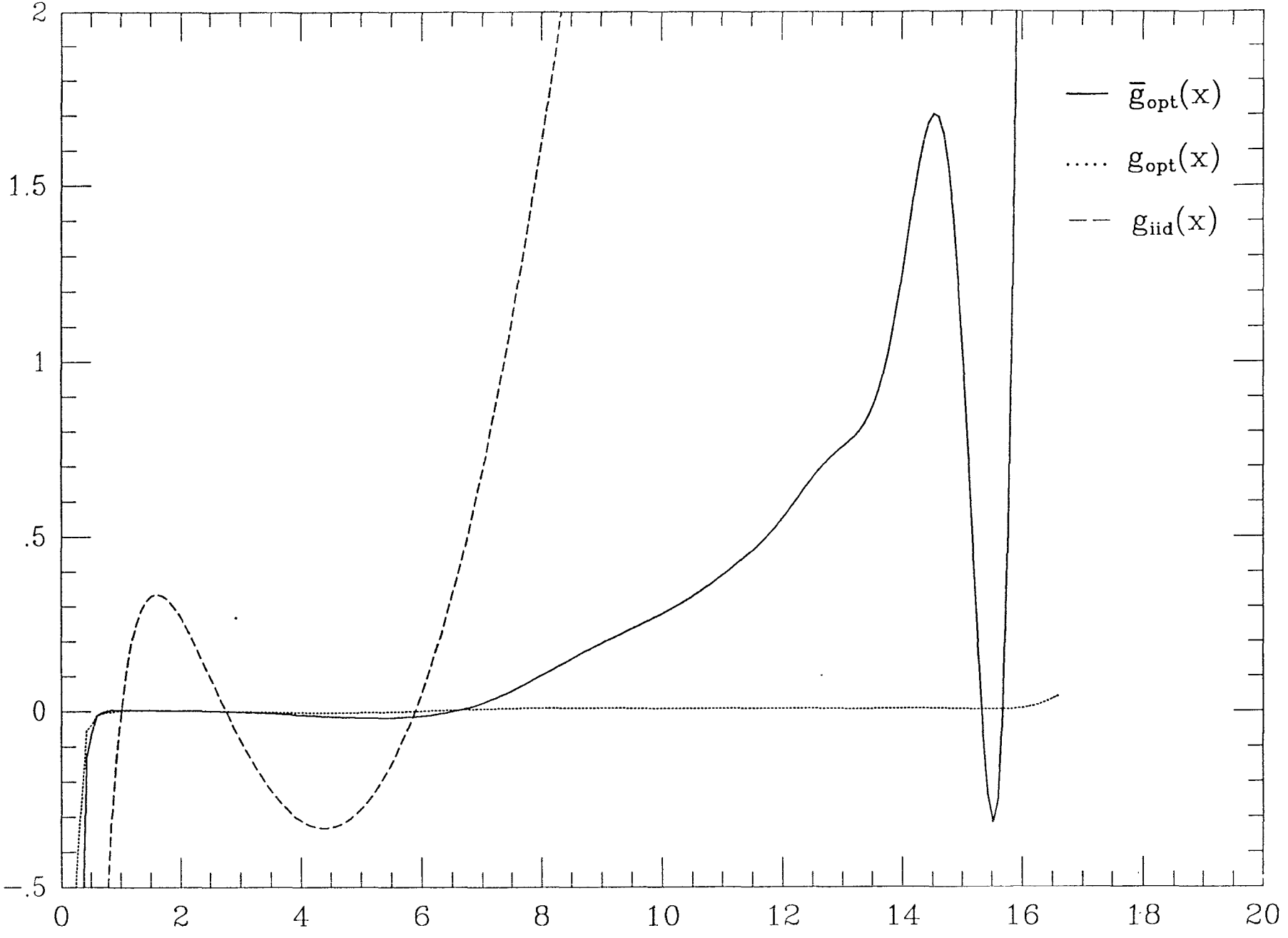


Fig. 5a Optimal Nonlinearities $\bar{g}_{opt}$ vs. g_{opt} and g_{iid} in Example 5

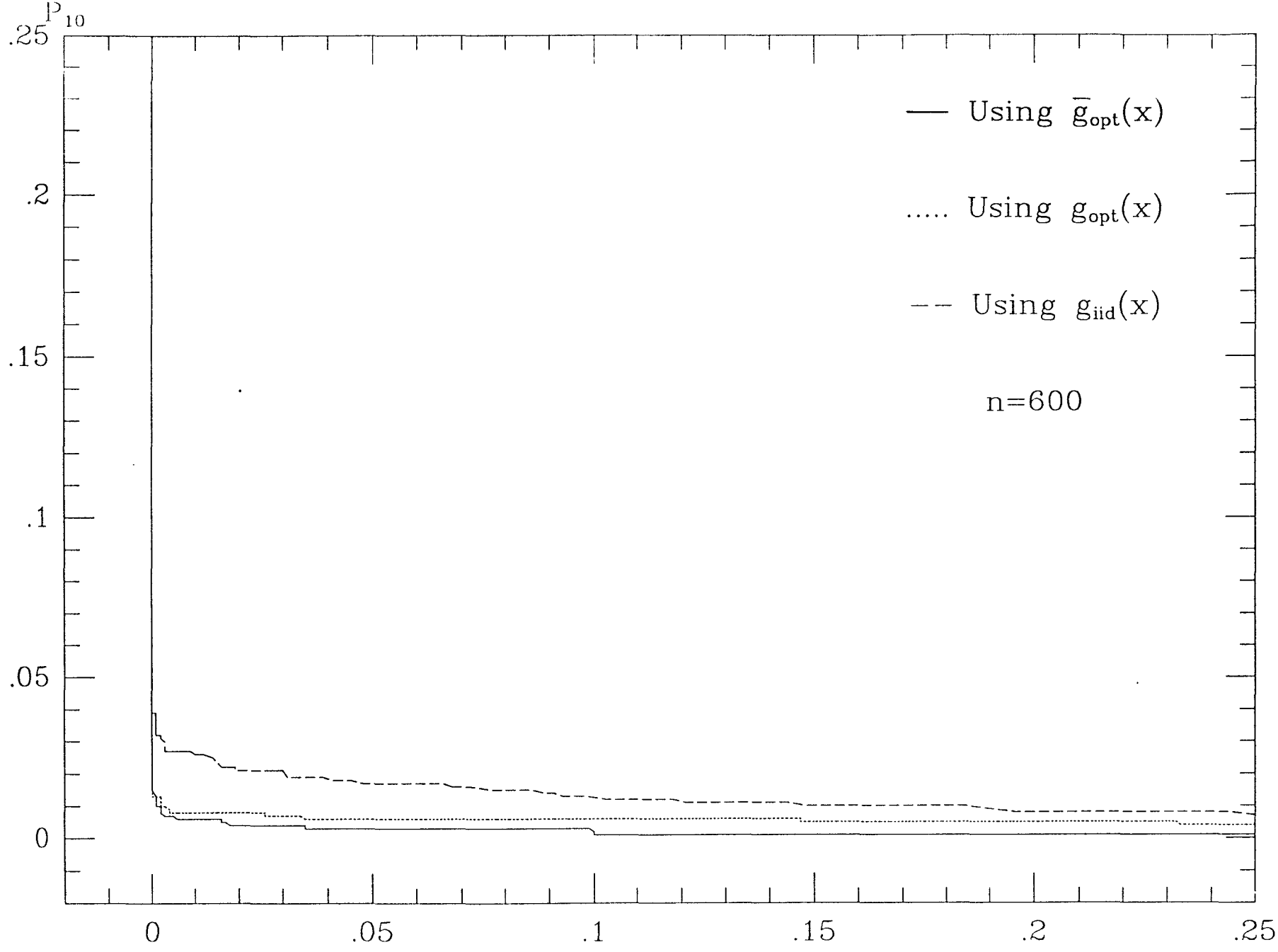


Fig. 5b Comparison of ROCs in Example 5

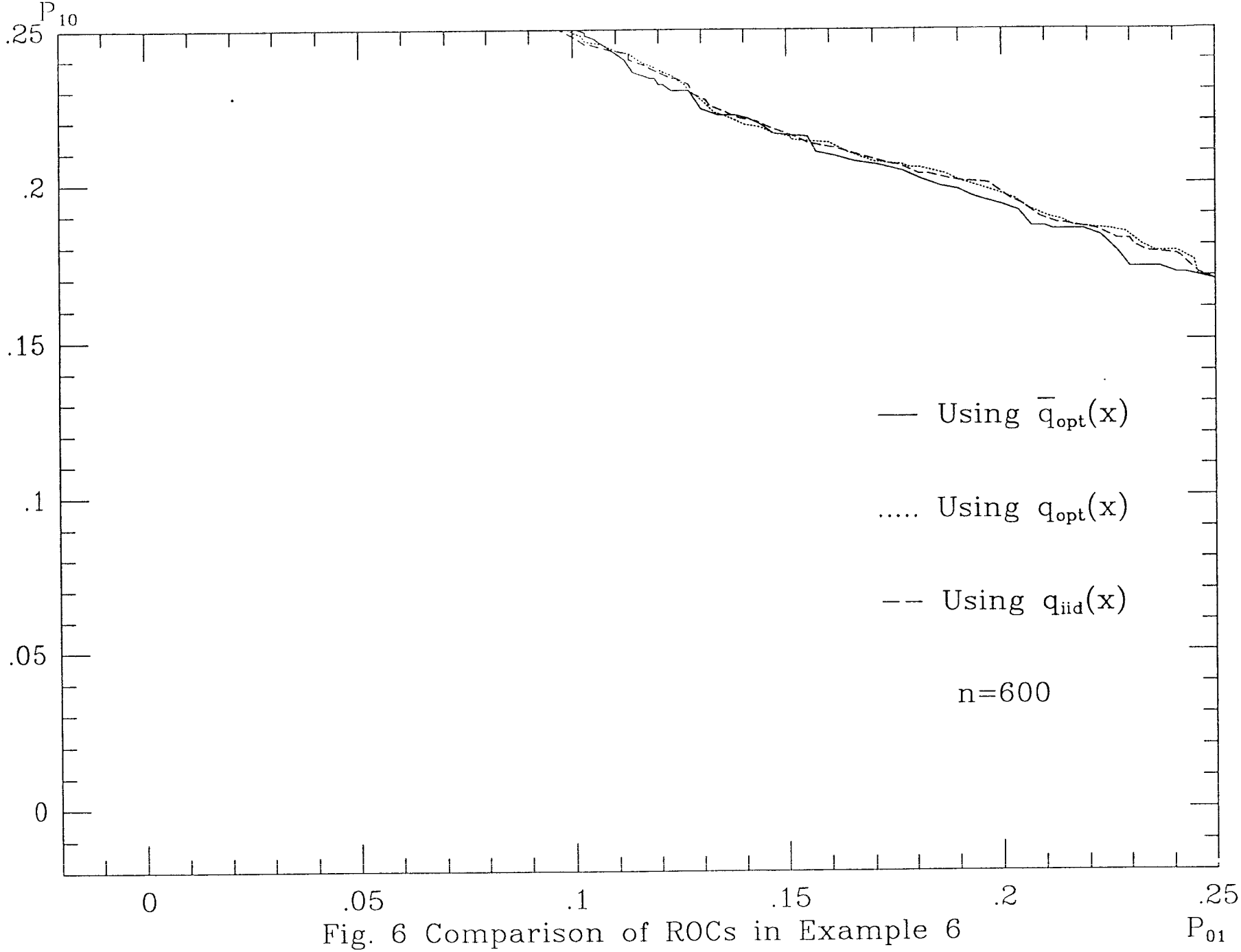


Fig. 6 Comparison of ROCs in Example 6

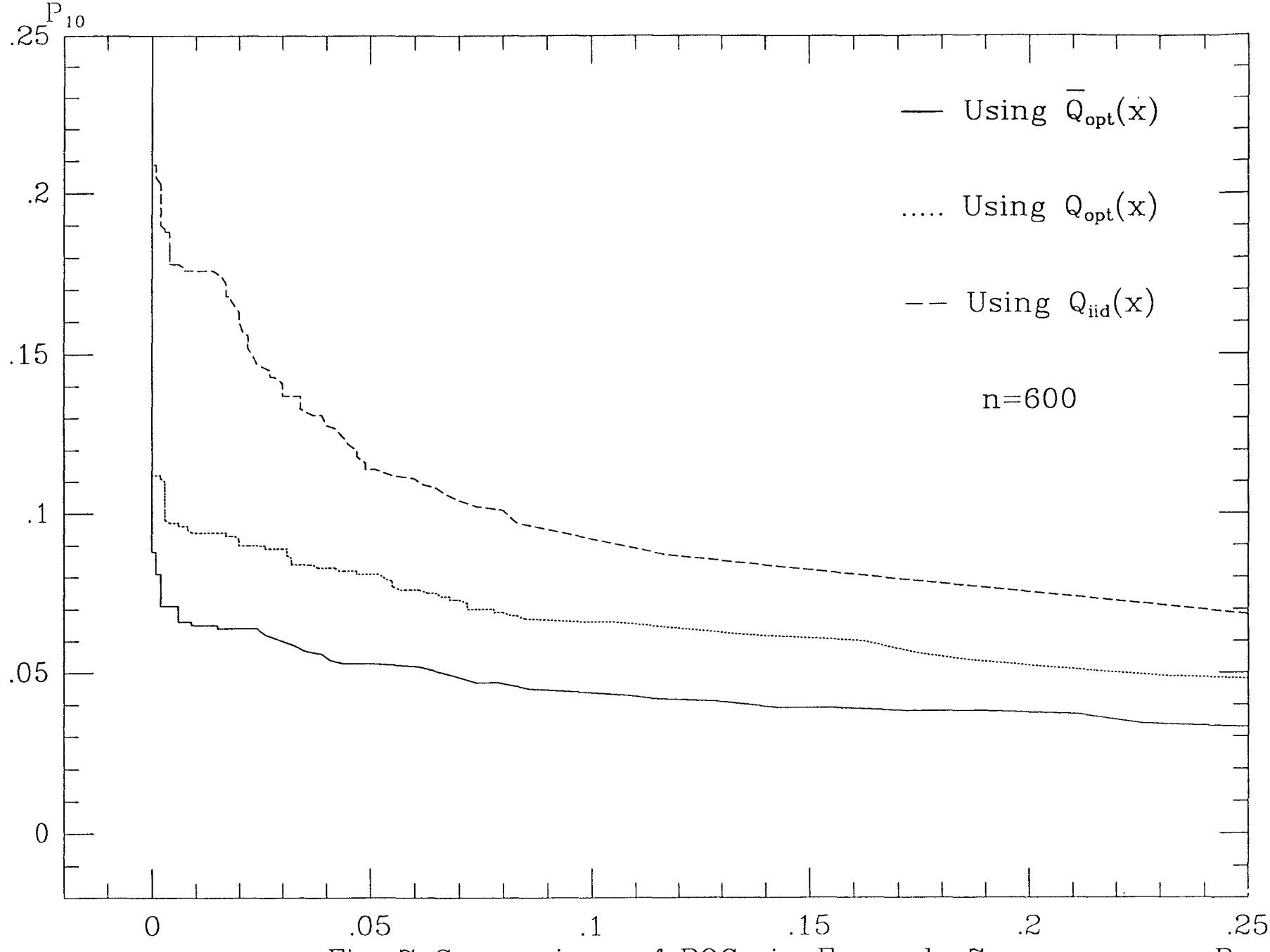


Fig. 7 Comparison of ROCs in Example 7

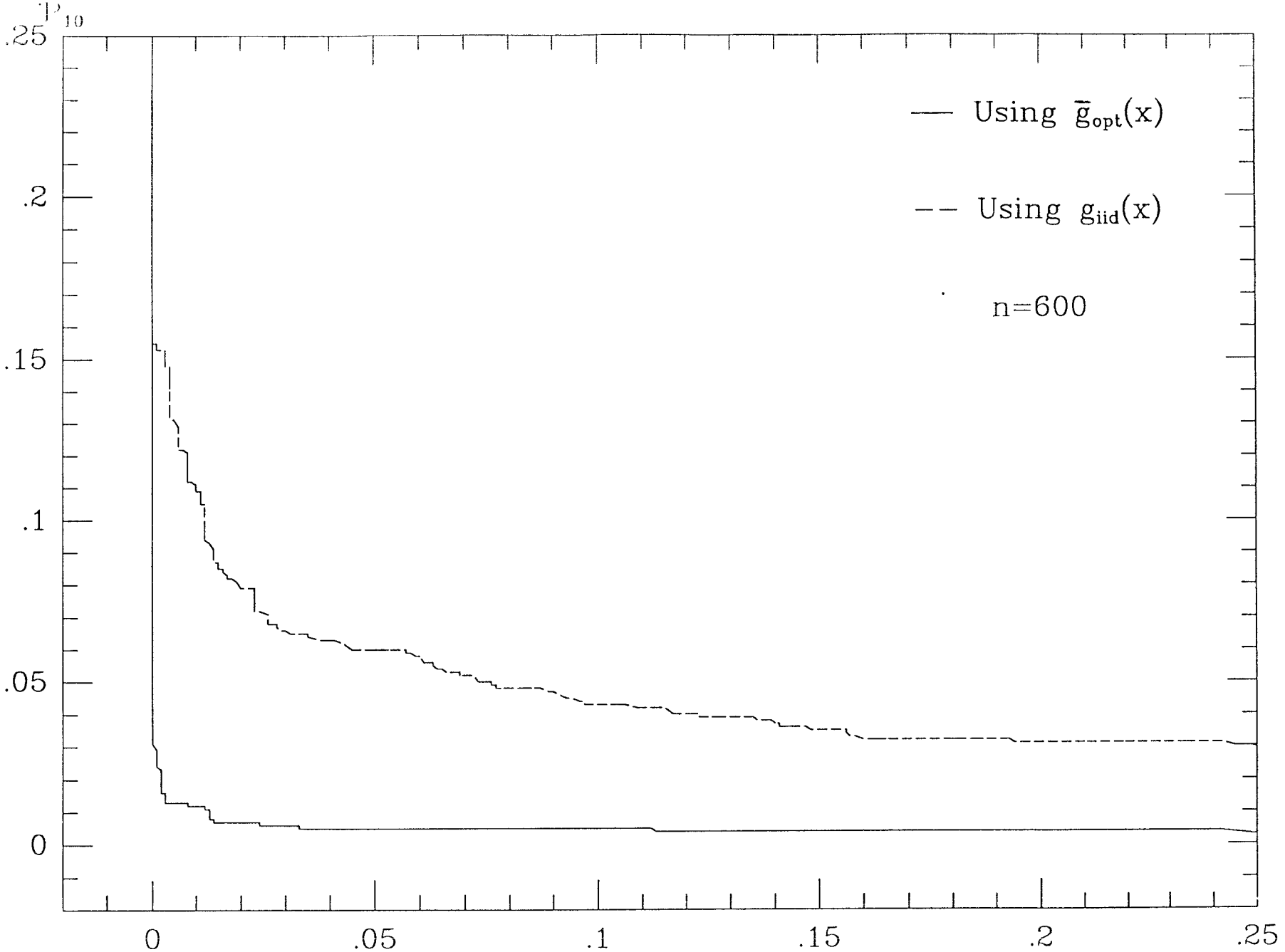


Fig. 8 Comparison of ROCs in Example 8

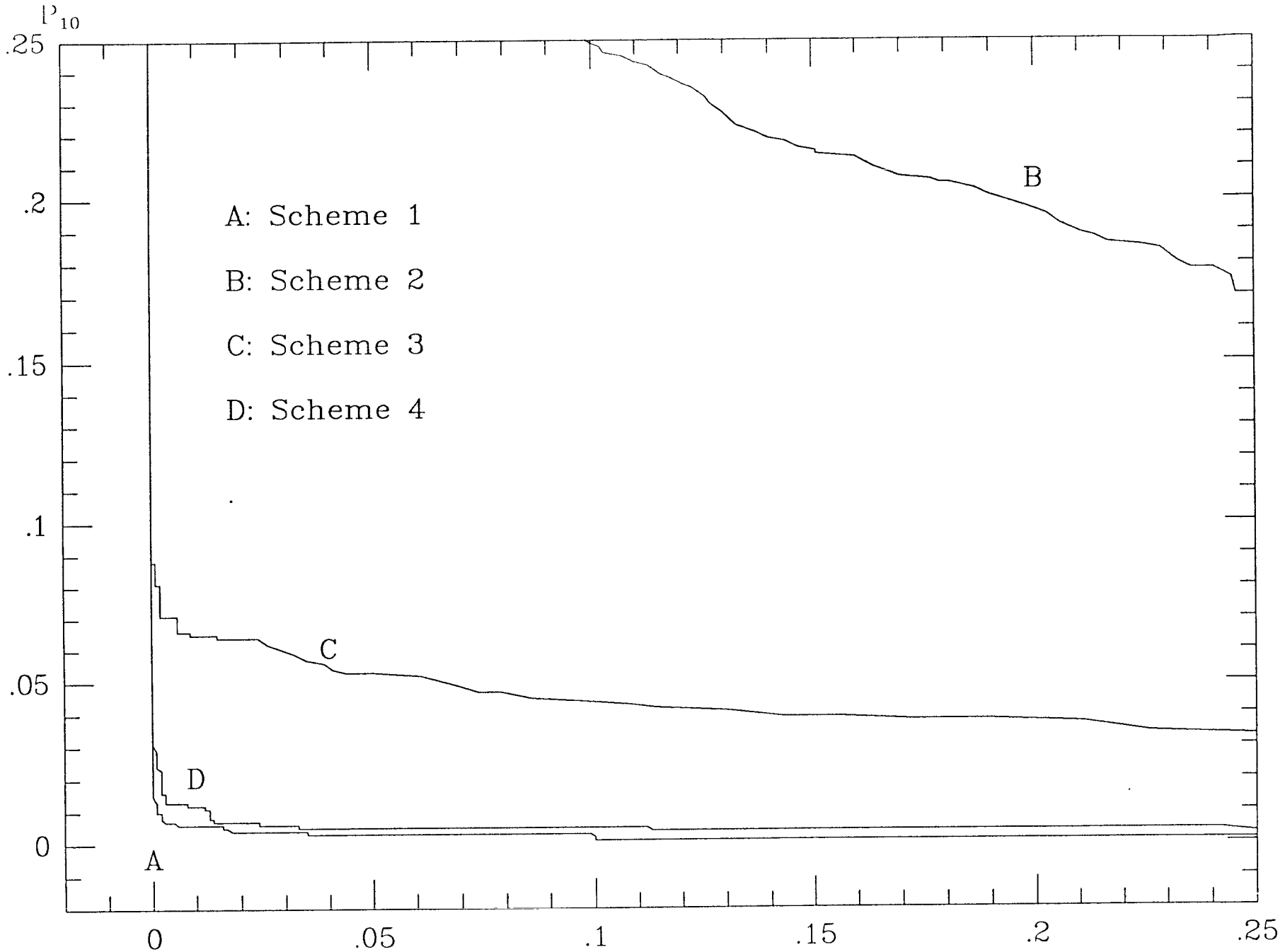


Fig. 8a Comparison of ROCs for Schemes 1–4 in the Case of Dependence Across Time and Sensors

q_{opt}: induced from g_{opt}		q_{iid}: induced from $g_{iid} = \ln(f_1/f_0)$	
breakpoints	*levels*	*breakpoints*	*levels*
0.02	-0.4546	0.02	-18.3392
2.0925	-0.1320×10^{-2}	2.0925	-0.4213×10^{-1}
4.1650	-0.1223×10^{-2}	4.1650	-0.7578×10^{-1}
6.2375	0.5736×10^{-2}	6.2375	1.0728
8.3100	0.7971×10^{-2}	8.3100	3.5113
10.3825	0.6954×10^{-2}	10.3825	7.2116
12.4550	0.6759×10^{-2}	12.4550	12.1400
14.5275	0.1441×10^{-1}	14.5275	18.2693

Table 1: 8-level Suoptimal Quantizers in Example 2

Q_{opt}		Q_{iid}	
breakpoints	*levels*	*breckpoints*	*levels*
0.02	0.9894	0.02	0.5766
0.2488	0.8298	0.4398	-0.3341
0.5059	-0.2231	0.5889	-0.4371×10^{-1}
0.7631	0.1629	0.7384	0.2217
1.0148	0.4922	0.9138	0.4045
1.2643	0.4816	1.0867	0.5478
1.5024	0.5263	1.5444	0.5317
1.7387	0.3725	2.6292	0.3401

Table 2: 8-level Optimal Quantizers in Example 3

$\bar{q}_{opt}$: induced from $\bar{g}_{opt}$	
breakpoints	*levels*
0.02	-6.0116
2.0925	-0.4211×10^{-2}
4.1650	-0.7475×10^{-2}
6.2375	0.3335×10^{-1}
8.3100	0.8737×10^{-1}
10.3825	0.2609
12.4550	0.3970
14.5275	3.9601

Table 3: 8-level Suboptimal Optimal Quantizer in Example 6

$\overline{Q}_{opt}$	
breakpoints	*levels*
0.02	-0.9918
0.2492	-0.8367
0.5056	-0.2216
0.7631	0.1662
1.0141	0.5000
1.2633	0.4685
1.5021	0.5061
1.7352	0.4063

Table 4: 8-level Optimal Quantizer in Example 7